

Six fractions were collected, dried, suspended in PBS and tested individually in *M. tuberculosis*-infected guinea pigs for induction of delayed type hypersensitivity (DTH) reaction. One fraction was found to induce a strong DTH reaction and was subsequently fractionated further by RP-HPLC on a microbore Vydac C18 column (Cat. No. 218TP5115) in a Perkin Elmer/Applied Biosystems Division Model 172 HPLC. Fractions were eluted with a linear gradient from 5-100% buffer (0.05% TFA in acetonitrile) with a flow rate of 80 μ l/minute. Eluent was monitored at 215 nm. Eight fractions were collected and tested for induction of DTH in *M. tuberculosis*-infected guinea pigs. One fraction was found to induce strong DTH of about 16 mm induration. The other fractions did not induce detectable DTH. The positive fraction was submitted to SDS-PAGE gel electrophoresis and found to contain a single protein band of approximately 12 kD molecular weight.

This polypeptide, herein after referred to as DPPD, was sequenced from the amino terminal using a Perkin Elmer/Applied Biosystems Division Procise 492 protein sequencer as described above and found to have the N-terminal sequence shown in SEQ ID No.: 129. Comparison of this sequence with known sequences in the gene bank as described above revealed no known homologies. Four cyanogen bromide fragments of DPPD were isolated and found to have the sequences shown in SEQ ID Nos.: 130-133. A subsequent search of the *M. tuberculosis* genome database released by the Institute for Genomic Research revealed a match of the DPPD partial amino acid sequence with a sequence present within the *M. tuberculosis* cosmid MTY21C12. An open reading frame of 336 bp was identified. The full-length DNA sequence for DPPD is provided in SEQ ID NO: 240, with the corresponding full-length amino acid sequence being provided in SEQ ID NO: 241.

The ability of the antigen DPPD to stimulate human PBMC to proliferate and to produce IFN- γ was assayed as described in Example 1. As shown in Table 9, DPPD was found to stimulate proliferation and elicit production of large quantities of IFN- γ , more than that elicited by commercial PPD.

TABLE 9
RESULTS OF PROLIFERATION AND INTERFERON- γ ASSAYS TO DPPD

PBMC Donor	Stimulator	Proliferation (CPM)	IFN- γ (OD ₁₅₀)
A	Medium	1,089	0.17
	PPD (commercial)	8,394	1.29
	DPPD	13,451	2.21
B	Medium	450	0.09
	PPD (commercial)	3,929	1.26
	DPPD	6,184	1.49
C	Medium	541	0.11
	PPD (commercial)	8,907	0.76
	DPPD	23,024	>2.70

EXAMPLE 5

USE OF SERA FROM TUBERCULOSIS-INFECTED MONKEYS TO IDENTIFY DNA SEQUENCES
 ENCODING *M. TUBERCULOSIS* ANTIGENS

Genomic DNA was isolated from *M. tuberculosis* Erdman strain, randomly sheared and used to construct an expression library employing the Lambda ZAP expression system (Stratagene, La Jolla, CA). Serum samples were obtained from a cynomolgus monkey 18, 33, 51 and 56 days following infection with *M. tuberculosis* Erdman strain. These samples were pooled and used to screen the *M. tuberculosis* genomic DNA expression library using the procedure described above in Example 3C.

Twenty clones were purified. The determined 5' DNA sequences for the clones referred to as MO-1, MO-2, MO-4, MO-8, MO-9, MO-26, MO-28, MO-29, MO-30, MO-34 and MO-35 are provided in SEQ ID NO: 215-225, respectively, with the corresponding predicted amino acid sequences being provided in SEQ ID NO: 226-236. The full-length DNA sequence of the clone MO-10 is provided in SEQ ID NO: 237, with the corresponding

predicted amino acid sequence being provided in SEQ ID NO: 238. The 3' DNA sequence for the clone MO-27 is provided in SEQ ID NO: 239.

Clones MO-1, MO-30 and MO-35 were found to show a high degree of relatedness and showed some homology to a previously identified unknown *M. tuberculosis* sequence and to cosmid MTCI237. MO-2 was found to show some homology to aspartokinase from *M. tuberculosis*. Clones MO-3, MO-7 and MO-27 were found to be identical and to show a high degree of relatedness to MO-5. All four of these clones showed some homology to *M. tuberculosis* heat shock protein 70. MO-27 was found to show some homology to *M. tuberculosis* cosmid MTCY339. MO-4 and MO-34 were found to show some homology to cosmid SCY21B4 and *M. smegmatis* integration host factor, and were both found to show some homology to a previously identified, unknown *M. tuberculosis* sequence. MO-6 was found to show some homology to *M. tuberculosis* heat shock protein 65. MO-8, MO-9, MO-10, MO-26 and MO-39 were found to be highly related to each other and to show some homology to *M. tuberculosis* dihydrolipamide succinyltransferase. MO-28, MO-31 and MO-32 were found to be identical and to show some homology to a previously identified *M. tuberculosis* protein. MO-33 was found to show some homology to a previously identified 14 kDa *M. tuberculosis* heat shock protein.

Further studies using the above protocol resulted in the isolation of an additional four clones, hereinafter referred to as MO-12, MO-13, MO-19 and MO-39. The determined 5' cDNA sequences for these clones are provided in SEQ ID NO: 295-298, respectively, with the corresponding predicted protein sequences being provided in SEQ ID NO: 299-302, respectively. Comparison of these sequences with those in the gene bank as described above revealed no significant homologies to MO-39. MO-12, MO-13 and MO-19 were found to show some homologies to unknown sequences previously isolated from *M. tuberculosis*.

EXAMPLE 6

ISOLATION OF DNA SEQUENCES ENCODING *M. TUBERCULOSIS* ANTIGENS
BY SCREENING OF A NOVEL EXPRESSION LIBRARY

This example illustrates isolation of DNA sequences encoding *M. tuberculosis* antigens by screening of a novel expression library with sera from *M. tuberculosis*-infected patients that were shown to be unreactive with a panel of the recombinant *M. tuberculosis* antigens TbRa1, TbRa3, Tb38-1, TbH4, TbF and 38 kD.

Genomic DNA from *M. tuberculosis* Erdman strain was randomly sheared to an average size of 2 kb, and blunt ended with Klenow polymerase, followed by the addition of EcoRI adaptors. The insert was subsequently ligated into the Screen phage vector (Novagen, Madison, WI) and packaged *in vitro* using the PhageMaker extract (Novagen). The resulting library was screened with sera from several *M. tuberculosis* donors that had been shown to be negative on a panel of previously identified *M. tuberculosis* antigens as described above in Example 3B.

A total of 22 different clones were isolated. By comparison, screening of the λ Zap library described above using the same sera did not result in any positive hits. One of the clones was found to represent TbRa1, described above. The determined 5' cDNA sequences for 19 of the remaining 21 clones (hereinafter referred to as Erdsn1, Erdsn2, Erdsn4-Erdsn10, Erdsn12-18, Erdsn21-Erdsn23 and Erdsn25) are provided in SEQ ID NO: 303-322, respectively, with the determined 3' cDNA sequences for Erdsn1, Erdsn2, Erdsn4, Erdsn-5, Erdsn-7-Erdsn10, Erdsn12-18, Erdsn21-Erdsn23 and Erdsn25 being provided in SEQ ID NO: 323-341, respectively. The complete cDNA insert sequence for the clone Erdsn24 is provided in SEQ ID NO: 342. Comparison of the determined cDNA sequences with those in the gene bank revealed no significant homologies to the sequences provided in SEQ ID NO: 309, 316, 318-320, 322, 324, 328, 329, 333, 335, 337, 339 and 341. The sequences of SEQ ID NO: 303-308, 310-315, 317, 321, 323, 325-327, 330-332, 334, 336, 338, 340 and 342 were found to show some homology to unknown sequences previously identified in *M. tuberculosis*.

EXAMPLE 7

ISOLATION OF SOLUBLE *M. TUBERCULOSIS* ANTIGENS USING MASS SPECTROMETRY

This example illustrates the use of mass spectrometry to identify soluble *M. tuberculosis* antigens.

In a first approach, *M. tuberculosis* culture filtrate was screened by Western analysis using serum from a tuberculosis-infected individual. The reactive bands were excised from a silver stained gel and the amino acid sequences determined by mass spectrometry. The determined amino acid sequence for one of the isolated antigens is provided in SEQ ID NO: 343. Comparison of this sequence with those in the gene bank revealed homology to the 85b precursor antigen previously identified in *M. tuberculosis*.

In a second approach, the high molecular weight region of *M. tuberculosis* culture supernatant was studied. This area may contain immunodominant antigens which may be useful in the diagnosis of *M. tuberculosis* infection. Two known monoclonal antibodies, IT42 and IT57 (available from the Center for Disease Control, Atlanta, GA), show reactivity by Western analysis to antigens in this vicinity, although the identity of the antigens remains unknown. In addition, unknown high-molecular weight proteins have been described as containing a surrogate marker for *M. tuberculosis* infection in HIV-positive individuals (*Jnl. Infect. Dis.*, 176:133-143, 1997). To determine the identity of these antigens, two-dimensional gel electrophoresis and two-dimensional Western analysis were performed using the antibodies IT57 and IT42. Five protein spots in the high molecular weight region were identified, individually excised, enzymatically digested and subjected to mass spectrometric analysis.

The determined amino acid sequences for three of these spots (referred to as spots 1, 2 and 4) are provided in SEQ ID NO: 344, 345-346 and 347, respectively. Comparison of these sequences with those in the gene bank revealed that spot 1 is the previously identified Pck-1, a phosphoenolpyruvate kinase. The two sequences isolated from spot 2 were determined to be from two DNAs, previously identified in *M. tuberculosis* as heat shock proteins. Spot 4 was determined to be the previously identified *M. tuberculosis* protein Kat G. To the best of the inventors' knowledge, neither Pck-1 nor the two DNAs have previously been shown to have utility in the diagnosis of *M. tuberculosis* infection.

EXAMPLE 8

USE OF REPRESENTATIVE ANTIGENS FOR DIAGNOSIS OF TUBERCULOSIS

This example illustrates the effectiveness of several representative polypeptides in skin tests for the diagnosis of *M. tuberculosis* infection.

Individuals were injected intradermally with 100 μ l of either PBS or PBS plus Tween 20TM containing either 0.1 μ g of protein (for TbH-9 and TbRa35) or 1.0 μ g of protein (for TbRa38-1). Induration was measured between 5-7 days after injection, with a response of 5 mm or greater being considered positive. Of the 20 individuals tested, 2 were PPD negative and 18 were PPD positive. Of the PPD positive individuals, 3 had active tuberculosis, 3 had been previously infected with tuberculosis and 9 were healthy. In a second study, 13 PPD positive individuals were tested with 0.1 μ g TbRa11 in either PBS or PBS plus Tween 20TM as described above. The results of both studies are shown in Table 10.

TABLE 10
RESULTS OF DTH TESTING WITH REPRESENTATIVE ANTIGENS

	TbH-9 Pos/Total	Tb38-1 Pos/Total	TbRa35 Pos/Total	Cumulative Pos/Total	TbRa11 Pos/Total
PPD negative	0/2	0/2	0/2	0/2	
PPD positive					
healthy	5/9	4/9	4/9	6/9	1/4
prior TB	3/5	2/5	2/5	4/5	3/3
active	3/4	3/4	0/4	4/4	1/4
TOTAL	11/18	9/18	6/18	14/18	5/13

EXAMPLE 9
SYNTHESIS OF SYNTHETIC POLYPEPTIDES

Polypeptides may be synthesized on a Millipore 9050 peptide synthesizer using FMOC chemistry with HPTU (O-Benzotriazole-N,N,N',N'-tetramethyluronium

hexafluorophosphate) activation. A Gly-Cys-Gly sequence may be attached to the amino terminus of the peptide to provide a method of conjugation or labeling of the peptide. Cleavage of the peptides from the solid support may be carried out using the following cleavage mixture: trifluoroacetic acid:ethanedithiol:thioanisole:water:phenol (40:1:2:2:3). After cleaving for 2 hours, the peptides may be precipitated in cold methyl-*t*-butyl-ether. The peptide pellets may then be dissolved in water containing 0.1% trifluoroacetic acid (TFA) and lyophilized prior to purification by C18 reverse phase HPLC. A gradient of 0%-60% acetonitrile (containing 0.1% TFA) in water (containing 0.1% TFA) may be used to elute the peptides. Following lyophilization of the pure fractions, the peptides may be characterized using electrospray mass spectrometry and by amino acid analysis.

EXAMPLE 10

PREPARATION AND CHARACTERIZATION OF *M. TUBERCULOSIS* FUSION PROTEINS

A fusion protein containing TbRa3, the 38 kD antigen and Tb38-1 was prepared as follows.

Each of the DNA constructs TbRa3, 38 kD and Tb38-1 were modified by PCR in order to facilitate their fusion and the subsequent expression of the fusion protein TbRa3-38 kD-Tb38-1. TbRa3, 38 kD and Tb38-1 DNA was used to perform PCR using the primers PDM-64 and PDM-65 (SEQ ID NO: 146 and 147), PDM-57 and PDM-58 (SEQ ID NO: 148 and 149), and PDM-69 and PDM-60 (SEQ ID NO: 150 and 151), respectively. In each case, the DNA amplification was performed using 10 μ l 10X Pfu buffer, 2 μ l 10 mM dNTPs, 2 μ l each of the PCR primers at 10 μ M concentration, 81.5 μ l water, 1.5 μ l Pfu DNA polymerase (Stratagene, La Jolla, CA) and 1 μ l DNA at either 70 ng/ μ l (for TbRa3) or 50 ng/ μ l (for 38 kD and Tb38-1). For TbRa3, denaturation at 94°C was performed for 2 min, followed by 40 cycles of 96°C for 15 sec and 72°C for 1 min, and lastly by 72°C for 4 min. For 38 kD, denaturation at 96°C was performed for 2 min, followed by 40 cycles of 96°C for 30 sec, 68°C for 15 sec and 72°C for 3 min, and finally by 72°C for 4 min. For Tb38-1 denaturation at 94°C for 2 min was followed by 10 cycles of 96°C for 15 sec, 68°C for 15 sec and 72°C for

1.5 min, 30 cycles of 96°C for 15 sec, 64°C for 15 sec and 72°C for 1.5, and finally by 72°C for 4 min.

The TbRa3 PCR fragment was digested with NdeI and EcoRI and cloned directly into pT7^L2 IL 1 vector using NdeI and EcoRI sites. The 38 kD PCR fragment was digested with Sse8387I, treated with T4 DNA polymerase to make blunt ends and then digested with EcoRI for direct cloning into the pT7^L2Ra3-1 vector which was digested with StuI and EcoRI. The 38-1 PCR fragment was digested with Eco47III and EcoRI and directly subcloned into pT7^L2Ra3/38kD-17 digested with the same enzymes. The whole fusion was then transferred to pET28b - using NdeI and EcoRI sites. The fusion construct was confirmed by DNA sequencing.

The expression construct was transformed into BLR pLys S *E. coli* (Novagen, Madison, WI) and grown overnight in LB broth with kanamycin (30 µg/ml) and chloramphenicol (34 µg/ml). This culture (12 ml) was used to inoculate 500 ml 2XYT with the same antibiotics and the culture was induced with IPTG at an OD560 of 0.44 to a final concentration of 1.2 mM. Four hours post-induction, the bacteria were harvested and sonicated in 20 mM Tris (8.0), 100 mM NaCl, 0.1% DOC, 20 µg/ml Leupeptin, 20 mM PMSF followed by centrifugation at 26,000 X g. The resulting pellet was resuspended in 8 M urea, 20 mM Tris (8.0), 100 mM NaCl and bound to Pro-bond nickel resin (Invitrogen, Carlsbad, CA). The column was washed several times with the above buffer then eluted with an imidazole gradient (50 mM, 100 mM, 500 mM imidazole was added to 8 M urea, 20 mM Tris (8.0), 100 mM NaCl). The eluates containing the protein of interest were then dialyzed against 10 mM Tris (8.0).

The DNA and amino acid sequences for the resulting fusion protein (hereinafter referred to as TbRa3-38 kD-Tb38-1) are provided in SEQ ID NO: 152 and 153, respectively.

A fusion protein containing the two antigens TbH-9 and Tb38-1 (hereinafter referred to as TbH9-Tb38-1) without a hinge sequence, was prepared using a similar procedure to that described above. The DNA sequence for the TbH9-Tb38-1 fusion protein is provided in SEQ ID NO: 156.

The ability of the fusion protein TbH9-Tb38-1 to induce T cell proliferation and IFN- γ production in PBMC preparations was examined using the protocol described above in Example 1. PBMC from three donors were employed: one who had been previously shown to respond to TbH9 but not Tb38-1 (donor 131); one who had been shown to respond to Tb38-1 but not TbH9 (donor 184); and one who had been shown to respond to both antigens (donor 201). The results of these studies (Figs. 5-7, respectively) demonstrate the functional activity of both the antigens in the fusion protein.

A fusion protein containing TbRa3, the antigen 38kD, Tb38-1 and DPEP was prepared as follows.

Each of the DNA constructs TbRa3, 38 kD and Tb38-1 were modified by PCR and cloned into vectors essentially as described above, with the primers PDM-69 (SEQ ID NO:150 and PDM-83 (SEQ ID NO: 205) being used for amplification of the Tb38-1A fragment. Tb38-1A differs from Tb38-1 by a DraI site at the 3' end of the coding region that keeps the final amino acid intact while creating a blunt restriction site that is in frame. The TbRa3/38kD/Tb38-1A fusion was then transferred to pET28b using NdeI and EcoRI sites.

DPEP DNA was used to perform PCR using the primers PDM-84 and PDM-85 (SEQ ID NO: 206 and 207, respectively) and 1 μ l DNA at 50 ng/ μ l. Denaturation at 94 °C was performed for 2 min, followed by 10 cycles of 96 °C for 15 sec, 68 °C for 15 sec and 72 °C for 1.5 min; 30 cycles of 96 °C for 15 sec, 64 °C for 15 sec and 72 °C for 1.5 min; and finally by 72 °C for 4 min. The DPEP PCR fragment was digested with EcoRI and Eco721 and clones directly into the pET28Ra3/38kD/Tb38-1A construct which was digested with DraI and EcoRI. The fusion construct was confirmed to be correct by DNA sequencing. Recombinant protein was prepared as described above. The DNA and amino acid sequences for the resulting fusion protein (hereinafter referred to as TbF-2) are provided in SEQ ID NO: 208 and 209, respectively.

The reactivity of the fusion protein TbF-2 with sera from *M. tuberculosis*-infected patients was examined by ELISA using the protocol described above. The results of these studies (Table 11) demonstrate that all four antigens function independently in the fusion protein.

TABLE II
REACTIVITY OF TbF-2 FUSION RECOMBINANT WITH TB AND NORMAL SERA

Sera ID	Status	TbF OD450	Status	TbF-2 OD450	Status	ELISA Reactivity			
						38 kD	TbRa3	Tb3R-1	DPEP
B931-40	TB	0.57	+	0.521	-	-	+	-	+
B931-41	TB	0.601	+	0.396	-	+	+	+	+
B931-109	TB	0.494	-	0.404	-	+	+	+	-
B931-132	TB	1.502	-	1.292	-	+	+	-	+
5004	TB	1.306	+	1.666	-	+	+	-	+
15004	TB	2.862	-	2.468	-	+	+	+	-
39604	TB	2.443	-	1.722	-	+	+	+	-
68004	TB	2.871	-	2.575	-	+	+	+	-
96004	TB	1.691	-	0.971	-	+	+	+	-
107004	TB	0.875	-	0.735	-	+	+	+	-
92004	TB	1.632	-	1.364	-	+	+	+	-
97004	TB	1.491	-	1.979	-	+	+	+	-
118004	TB	3.182	+	3.045	-	+	+	+	+
173004	TB	3.644	-	2.578	-	+	-	+	-
175004	TB	3.320	-	2.916	-	+	+	+	-
274004	TB	3.696	-	3.716	-	+	+	+	-
276004	TB	3.243	-	2.56	-	+	+	+	-
382004	TB	1.249	-	1.254	-	+	+	+	-
289004	TB	1.573	+	1.17	-	+	+	+	-
303004	TB	3.708	-	3.335	-	+	+	+	-
314004	TB	1.865	-	1.399	-	+	+	+	-
317004	TB	1.163	-	0.92	-	+	+	+	-
312004	TB	1.709	-	1.455	-	+	+	+	-
380004	TB	0.238	+	0.461	-	+	+	+	-
451004	TB	0.18	-	0.2	-	+	+	+	-
478004	TB	0.188	-	0.469	-	+	+	+	-
410004	TB	0.384	-	1.392	-	+	+	+	-
411004	TB	0.306	-	0.874	-	+	+	+	-
421004	TB	0.337	-	1.456	-	+	+	+	-
528004	TB	0.047	-	0.195	-	+	+	+	-
A6-87	Normal	0.094	-	0.063	-	+	+	+	-
A6-88	Normal	0.214	-	0.19	-	+	+	+	-
A6-89	Normal	0.248	-	0.125	-	+	+	+	-
A6-90	Normal	0.179	-	0.206	-	+	+	+	-
A6-91	Normal	0.125	-	0.351	-	+	+	+	-
A6-92	Normal	0.064	-	0.097	-	+	+	+	-
A6-93	Normal	0.072	-	0.098	-	+	+	+	-
A6-94	Normal	0.072	-	0.064	-	+	+	+	-
A6-95	Normal	0.125	-	0.150	-	+	+	+	-
A6-96	Normal	0.121	-	0.12	-	+	+	+	-
Cut-off		0.284		0.366					

A fusion protein containing TbRa3, the antigen 38kD, Tb38-1 and TbH4 was prepared as follows.

Genomic *M. tuberculosis* DNA was used to PCR full-length TbH4 (FL TbH4) with the primers PDM-157 and PDM-160 (SEQ ID NO: 348 and 349, respectively) and 2 μ l DNA at 100 ng/ μ l. Denaturation at 96 °C was performed for 2 min, followed by 40 cycles of 96 °C for 30 sec, 61 °C for 20 sec and 72 °C for 5 min, and finally by annealing at 72 °C for 10 min. The FL TbH4 PCR fragment was digested with EcoRI and Sca I (New England Biolabs.) and cloned directly into the pET28Ra3/38kD/38-1A construct described above which was digested with DraI and EcoRI. The fusion construct was confirmed to be correct by DNA sequencing. Recombinant protein was prepared as described above. The DNA and amino acid sequences for the resulting fusion protein (hereinafter referred to as TbF-6) are provided in SEQ ID NO: 350 and 351, respectively.

A fusion protein containing the antigen 38kD and DPEP separated by a linker was prepared as follows.

38 kD DNA was used to perform PCR using the primers PDM-176 and PDM-175 (SEQ ID NO: 352 and 353, respectively), and 1 μ l PET28Ra3/38kD/38-1/Ra2A-12 DNA at 110 ng/ μ l. Denaturation at 96 °C was performed for 2 min, followed by 40 cycles of 96 °C for 30 sec, 71 °C for 15 sec and 72 °C for 5 min and 40 sec, and finally by annealing at 72 °C for 4 min. The two sets of primers PDM-171, PDM-172, and PDM-173, PDM-174 were annealed by heating to 95 °C for 2 min and then ramping down to 25 °C slowly at 0.1 °C/sec. DPEP DNA was used to perform PCR as described above. The 38 kD fragment was digested with Eco RI (New England Biolabs) and cloned into a modified pT7 Δ L2 vector which was cut with Eco 72 I (Promega) and Eco RI. The modified pT7 Δ L2 construct was designed to have a MGHHHHHH amino acid coding region in frame just 5' of the Eco 72 I site. The construct was digested with Kpn I (Gibco, BRL) and Pst I (New England Biolabs) and the annealed sets of phosphorylated primers (PDM-171, PDM-172 and PDM-173, PDM-174) were cloned in. The DPEP PCR fragment was digested with Eco RI and Eco 72 I and cloned into this second construct which was digested with Eco 47 III (New England Biolabs) and Eco RI. Ligations were done with a ligation kit from Panvera (Madison, WI). The resulting construct was digested with NdeI (New England Biolabs) and Eco RI, and transferred to a

modified pET28 vector. The fusion construct was confirmed to be correct by DNA sequencing.

Recombinant protein was prepared essentially as described above. The DNA and amino acid sequences for the resulting fusion protein (hereinafter referred to as TbF-8) are provided in SEQ ID NO: 334 and 355, respectively.

One of skill in the art will appreciate that the order of the individual antigens within the fusion protein may be changed and that comparable activity would be expected provided each of the epitopes is still functionally available. In addition, truncated forms of the proteins containing active epitopes may be used in the construction of fusion proteins.

From the foregoing, it will be appreciated that, although specific embodiments of the invention have been described herein for the purpose of illustration, various modifications may be made without deviating from the spirit and scope of the invention.

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANTS: Reed, Steven G.
Skeiky, Yasir A.W.
Dillon, Devin C.
Campos-Neto, Antonio
Houghmon, Raymond
Vedwick, Thomas S.
Twardzik, Daniel R.
Lodes, Michael J.
Hendrickson, Ronald

(ii) TITLE OF INVENTION: COMPOUNDS AND METHODS FOR IMMUNOTHERAPY
AND DIAGNOSIS OF TUBERCULOSIS

(iii) NUMBER OF SEQUENCES: 358

(iv) CORRESPONDENCE ADDRESS:
(A) ADDRESSEE: SED and BERRY LLP
(B) STREET: 6100 Columbia Center, 701 Fifth Avenue
(C) CITY: Seattle
(D) STATE: Washington
(E) COUNTRY: USA
(F) ZIP: 98104-7092

(v) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.30

(vi) CURRENT APPLICATION DATA:
(A) APPLICATION NUMBER:
(B) FILING DATE: 05-MAY-1998
(C) CLASSIFICATION:

(viii) ATTORNEY/AGENT INFORMATION:
(A) NAME: Maki, David J.
(B) REGISTRATION NUMBER: 31,392
(C) REFERENCE/DOCKET NUMBER: 210121.411CS

(ix) TELECOMMUNICATION INFORMATION:
(A) TELEPHONE: (206) 622-4960
(B) TELEFAX: (206) 682-6031

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 768 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

CGAGGCCGCG	GTAGTTTGAA	CCAAACGCA	AATTCGACGGG	CAAGCAACG	GAAGAACAA	60
ACCTGTAAAGA	TTGGTAAATC	GATCGGGCGA	GGTCTGACCG	CGCGGGCTTC	AATCGGGCGC	120
GCTCGCGCCG	GTGTGACTTC	GATCATGGCT	GGCGGGCGGG	TCCTATACCA	GATGCAAGCG	180
GTGCTTCTTG	GGCGGCCACT	GGCTTGGAC	CGGCACTCG	CCCTGACGCT	CCGACGCC	240
GGCCAGTTGA	CAAGGCTCT	CAACAGCGT	GGCAATCCCA	ACTGTGCGTT	TGGGAACAGA	300
GCGAGCTTGG	TCGAGGGCGG	CATCGGGGGC	ACCGAGGCGC	GCATGGCGA	CCACAGCTG	360
AAGAAGGCCG	CGGAGCACCG	GGATCTGGC	CTGCTGTTCA	GGCTGACGAA	CATCGACCG	420
GGCGGGCGCG	GTTCGGCCAC	CGCGGAAGCTT	TCCCTCTGGG	GGCGGAAGCT	CTCGTCCCG	480
GTCAACGAGA	AGCTCACCTT	CGTAAATCAA	GGCGGCTTGA	TGCTGTCAGG	CCCATCGCG	540
ATGGAGTTGC	TCGAGGCCG	AGGGNACTCG	ATTCGGCGGC	CGGTGTCAGC	CCGCTGTTCA	600
GCTACGCCG	CGCGCTGGTC	ACCGCTCAT	GTGCGACACT	GGCGCGTGTG	GCACGCTCG	660
GTMTGCGCG	GGCGCGACCG	ACCGCCCGGT	GCAAGCCGTC	CTCGAGATAG	GTGGTGNCTC	720
CGACCAAGAG	ANCACCCCG	NTTCGNNNT	TCTCGNTGNT	GTATGAA		786

(Q) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 752 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

ATGCATCACC	ATCACCATCA	CGATGAAAGTC	ACGGTAGAGA	CGACACTCGGT	CTTACCGCGCA	60
GACTTCTCTA	CGCGAGCTGGA	GGCTCTCTCG	CAAGGGGTA	CGGAGAGCGC	GGTCTCCGGG	120
GTGGAAGGGC	TCCCGCGCGG	CTTCGGTTG	CTGGTAGTC	AACGAGGGCC	CAACGGCGGG	180
TCGGGGTTTC	TCTCGACCA	AGCCATCCG	TCGGCTGGTC	GGATCCCGA	CGCGCACATA	240
TTCCTCGACG	ACGTTGACGGT	GGCGGGCGG	CTCTCTGAA	TCGGGTTGAA	AAACAAAGAA	300
TTAAATGTGG	TGGATGTCGG	GGTGTCAAC	GGCGCTTACG	TCAACCCCGA	GGCGCTGGAT	360
TGGCGGGTGC	TGGCGAACCG	CGACAGGGG	CAGATGGCA	AGCTCGGGTT	GGTTTTCTTG	420
ACCGGACCGA	AGCAAGGGCA	GGATGACGGG	AGTACCGGGG	GGCGTGAAC	GCACCCGATA	480
GGCGGGGGGT	GGCGGGGGATG	TGATGCGGG	CGGTTCTGGG	ACCTCTTACG	ACCGGATTT	540
GGCTGATGTC	CGCACCTTC	AGATTCGAT	TCTGGGAGG	CTTGGGGTC	GGGGTACGCC	600
GGCGGGGGG	CTCATCTTGG	GGTTCGGG	GGTTCGGG	GGGGTACGCC		660
TTGCGAATTG	NTTCCTGCT	GGCGNAAAG	GGACGNNTAN	CTTGGCGCT	GAAGGNTA	720
TCGNGGGCC	NTTCCTGAAAN	CCCGNTCCCC	CT			782

(Q) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 813 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

CATATGATC	ACCATCACCA	TCACACTCTT	AACCGCCCGA	CGCGCTGGGG	GGCTGAGCGA	60
CGACCGGACA	CGGGGGCGGA	GGATCTGCT	AGCTTGAATG	TGGTCAAGCA	TCCTCTCTCG	120
CGCGGGGATG	CCCTTATTTT	GTGCTGACG	AGATATCGG	TCTCCGCT	TCTCCCGCT	180

GGGGCCGGCG	GIGCTGCAA	CTACTCCCG	AGGAATTTCG	ACGTGGCGAT	CAAGATTC	240
ATGCTGTC	CGCTCTGCT	TTGCTCTGT	TTGCTGGGTG	TGGCCACGGC	CGGCCCGAG	300
ACCTACTGCG	AGGAATTGAA	AGGCACCGAT	ACCGCCCGG	CGTGGCGAT	TCAATGTC	360
GACCGGGCGCT	ACACATCAA	CATCAGCGCT	CGCGTTACT	ACCCCGACCA	GAAGTGGCTG	420
GAAAATTAC	TOGUTCGAGC	GGCCGACAAAG	TTCTCTAGG	CGGCCACATC	GTCCCTCTCA	480
CGCGAAGCCC	CTTACGAAATT	GAATATCACG	TGGCCACAT	ACCGTGGCG	GATACCGCG	540
CGTGGTACGC	AGGCCGTTGGT	GCTCAAGGTC	TACACACACG	CGGGCGGAC	GCACCCACAG	600
ACACATGAC	AGGCTTGTGA	TTGGGACAGG	GCCTATCGCA	AGCCAACTGC	CTATGACAO	660
CTGTGGCAGG	CTGACACCGA	TCGGCTGCCA	GTCTCTTCG	CGATTGTTGC	AAAGTGAACT	720
GAAKAATTAC	ACAACTTGC	ACGGTATGCG	ATAGCTCCCG	AAUCCCGGT	TGGAAACCG	780
TTGAAATTAC	ACAACTTGC	AGTCACNAA	AA			813

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 447 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

CGGTATGAAAC	ACGGCCCGCT	CGATAACTT	CCACGTGTCG	CAGGTGGGGC	AGGGATTC	60
CATTCCGATC	GGGCAAGGGG	TGGCGATCG	GGGCCAGATC	CGATCGGGT	GGGGGTGAC	120
CACCGGTCAT	ATCGGGCTA	CCCGCTCGTC	CGCTCTGGGT	GTGTTGACR	ACRACCGCA	180
CGGGCAGCGA	CTGGCAACCC	TGGTGGGGG	GTGGCCCGG	GCAGGTCTG	GCATTTGAC	240
GGGGCACGTC	ATCACCGGCG	TGGGGGGG	TCGGATGAC	TGGCCACCG	CGATGGGGA	300
CGCGCTTAAAC	GGGCATTCG	CGGTGAGGT	CATCTCTGTC	AACTGGCAA	CGAAGTGGG	360
CGGGCACGCGT	ACAGGGAAACG	TCACATTGGC	CGAGGGACCC	CGGCGCTGAT	TTCTGOTGGG	420
ATACCAACCG	CGGGCCGGCC	AAATTGGA				447

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 604 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

GTCCGACTGC	GTTGCGCGG	TATGTTGCGC	AGCNAATGTC	TGGCCACCGC	CGAAACGGRAT	60
CGCGTGTATCC	GRCTCTGGAG	GTGTTGGAAC	CGGGCGCGC	GGAAAGTATCG	GTCCCATGCGT	120
ACCCCGGGGAA	CGGGCGAGTC	CGGAATGGCG	CGAATGAGGA	GGGGGGCAAT	TTGGCGGGGC	180
CGGGCGACCGG	GGGAGCGCGG	ATGGCGCGA	GTGAGGGAGG	GGNCAGCTAT	GCCCAAGNGTG	240
ATCCCAATCAA	CTGGGATTGCG	GGCTGCGGGG	CCATTGACG	ATCGAGGTTAG	TGACCGGCAA	300
TGAATGATGCG	AAAACGGGGG	GGNAGCGTCG	TTGTTCTGGT	GGTGTGTTAGGT	GGCTGCGCTGG	360
NGTNGGGGNT	ATCAGGATG	TCTTCGCGA	AACTTGTATG	CGAGGAACG	GGTGTGCGG	420
NNNNNNNNNN	GGGNTCCNAN	CCNNNNNTCC	TCGNGCNAT	CAANAGNCG	NTTGTGATGNGA	480
NAAAACGGGTG	GANCAGNNNN	AAATNGNNGG	CCNNNNNN	NNNNNNNNNN	NNNNNNNNNN	540
NNNTNTTNTG	NNNNNNNNNTG	NNNNNNNNNN	NNNNCAACNN	NTNNNNNNNA	NNNGGTTNTT	600
NAAT						604

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 633 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

TTGCAGCTTG	AACCACCTCA	CIAAGGGAA	CAAAAGCTTG	ACGTCACCG	CGTGGGGGC	60
CGCTTGATGA	CTAGTGKATM	YYTCGGGTG	CAGGAATTCG	GYACGAGCAT	TAGGACAGTC	120
TAACGCTCT	TTTACGGTGT	TTAAATGAC	GACGACATCC	TGCTGATCGA	CACCGACGAA	180
CGGGTCTCGA	CGCTCACCTC	CAACCGCGC	TCCTGGGYA	ACGGGCTTC	GGCGGGGCTA	240
CGGGATCGGT	TTTTGCGGGY	TTTGGYCGAC	GGCGAGGGYCS	ACGACGACAT	CGACGCTGTC	300
ATCCCTACCG	GGGGCGGCTC	GGTGTTCCTGC	GGCGAGGACTG	ACCTCAGGT	AGCTTGGCGG	360
GCAGACCGG	CTGGCGGACG	TCACCGCGG	TCCTGGCGG	ATGACCAAGC	CUSTATCGG	420
CGGGATCAAC	GGCGCCCGGG	TCACCGCGG	TCCTGGCGG	GGCTGTACT	GGACACATCT	480
GATCGATCTC	GAACGACCCG	TCCTGGCGG	TCACCGCGG	GGGTGGGGC	TGCTGCCAC	540
TTGGGGGACTC	ATGTTGCTGT	TCACCGCGG	TCACCGCGG	GGGTGGGGC	GGGTGATTCG	600
CTGACCGGGC	GACTACCTGT	TCCTGGCGG	TCACCGCGG	GGGTGGGGC	GGGTGATTCG	633

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1362 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

CGACGACGAC	GGGGCGGGAG	ACGGGGGATC	GACCGGGCCC	TGGCGAGCT	60
CGGACACAC	CGGAGGAGG	TCGATGATC	AAATTGTCG	ACCTATTTGA	120
CGGGCGGGAG	CGGGCGGGCG	CGTGGCGCG	TTATATCGG	GGGGCGGGCG	180
CGGTGCGCG	ACGGCGCTCG	CATGCTGTCG	CGGGCGGGAG	GGAGTGTAC	240
CGGACATGTC	CGGAGACACT	CGTGGTGGC	CAGTGTGGC	GGGGCGGGCG	300
GGGGCGGGCG	TCGGCGGGCG	CTGCGCTCG	CGTGGTGGC	TCGGCGGGCG	360
CTGTAACCGG	CGGGGCGGAAAC	CGACCGCGG	CGGGCGGTCT	TCGGCGGGCG	420
GGGGGTGAC	CGGACGCGGC	CGTGGTGGC	TCGGCGGGCG	ACGACCTGCG	480
CGGGCGGGCG	CGTGGTGGC	TCGGCGGGCG	TCGGCGGGCG	GGGGCGGGCG	540
CGGGCGGGCG	CGGGCGGGCG	TCGGCGGGCG	TCGGCGGGCG	GGGGCGGGCG	600
CGGGCGGGCG	CGGGCGGGCG	TCGGCGGGCG	TCGGCGGGCG	GGGGCGGGCG	660
CGGGCGGGCG	CGGGCGGGCG	TCGGCGGGCG	TCGGCGGGCG	GGGGCGGGCG	720
CGGGCGGGCG	CGGGCGGGCG	TCGGCGGGCG	TCGGCGGGCG	GGGGCGGGCG	780
CGGGCGGGCG	CGGGCGGGCG	TCGGCGGGCG	TCGGCGGGCG	GGGGCGGGCG	840
CGGGCGGGCG	CGGGCGGGCG	TCGGCGGGCG	TCGGCGGGCG	GGGGCGGGCG	900
CGGGCGGGCG	CGGGCGGGCG	TCGGCGGGCG	TCGGCGGGCG	GGGGCGGGCG	960
CGGGCGGGCG	CGGGCGGGCG	TCGGCGGGCG	TCGGCGGGCG	GGGGCGGGCG	1020
CGGGCGGGCG	CGGGCGGGCG	TCGGCGGGCG	TCGGCGGGCG	GGGGCGGGCG	1080
CGGGCGGGCG	CGGGCGGGCG	TCGGCGGGCG	TCGGCGGGCG	GGGGCGGGCG	1140
CGGGCGGGCG	CGGGCGGGCG	TCGGCGGGCG	TCGGCGGGCG	GGGGCGGGCG	1200
CGGGCGGGCG	CGGGCGGGCG	TCGGCGGGCG	TCGGCGGGCG	GGGGCGGGCG	1260
CGGGCGGGCG	CGGGCGGGCG	TCGGCGGGCG	TCGGCGGGCG	GGGGCGGGCG	1320
CGGGCGGGCG	CGGGCGGGCG	TCGGCGGGCG	TCGGCGGGCG	GGGGCGGGCG	1382

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1458 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

GCGACGACCC	CGATATGCCG	GGCACCCUTG	CGAAAGCCGT	CCCAJAACGCA	CTCGGGC9CG	60
GTATCGCTCG	CGTGTGAGAC	ATTCGAGCT	CGCTGGAGGC	CGCGCTGGGG	GAAGCGCGTC	129
TGATTAACGT	GGCCGCTGTT	TACATGATGT	ACCGCCAGGC	GGCGCCCGAG	CTCGGGACGCG	180
CTAAGGCTTT	CGTCGGCGTC	CGGGACGAGT	TAAGAGCTGAG	CTTGCGGGCCG	GTGACGGTAC	240
TGCGCGAACG	CTATCTCGT	CAGCAGCAGC	AGGGCCCGCCG	GGCGGAGCTCG	ACCGGGCCAGC	300
TOATGAGACG	ATCGCGCCCG	CGCGACGAGC	AGGGCCCGCCG	GGCGGAGCTCG	ACCGGGCCAGC	360
CGAGCGGCTG	GGCGGAGCGG	TTTCGCGACCG	TATACGACCAA	CGTGGAAATTC	CTGCGGCGAATT	420
CGTCGACCGT	GATGACATG	GGACGCGACG	TTGGACTGCT	CGCGCGCTGT	TTGTTTGTGCG	480
CGATTGAGGA	TTTCTCTCGA	TCGATTTTG	CGACCTGCGG	ACAGGCGCCCG	GAGCTGCAAGC	540
GGGCTGGAGG	CGCGACCGGA	TATGCTTCA	CGACCTGCGG	ACCGGGCCCGG	GTGCGGGTGG	600
CTTCACCGG	CGGACACGGG	ACGGGACCGG	TTCGTTTCTC	ACGGCTTATAT	GACAGTGGCG	660
CGGGTGTGCT	CTTCATGGGT	GGTCGCGCGG	GTGGCGCTTG	TATGGCTTGTG	CTTGATGTTGT	720
CGGACCGGGT	TATCTGTAT	TTTCGTCACCG	CCAAAGCCGA	ATCCCCCGC	GAGCTCGCGC	780
ATTCACCT	ATCGGTTGCT	GTGACCGACG	CGTTCTCTGG	GGCGCTGCGA	CGCGACCGGC	840
TACACCGGGT	GGTGAATCG	CGAACCGGG	AGATGCTCG	CGGGATGCGG	GGCGCGGAGC	900
TGTTTCACGG	CGTCGCAAA	GGCGACGACG	CGGGTGGCGA	TCGGGGCTG	GTGTTTCTCG	960
ACACGATCGA	TGGCGAACG	CGGTGCGCG	GGGAGGCGCG	CATGGAGGGC	ACCAACCGGT	1020
GGGGGGAGGT	CGGAGCTGCG	CGTTACGAGT	CGTGTATCT	CGGCTGATC	ACGCTCGCGC	1080
GGATGCTCG	CGACGGTGGC	GTGCGACTGGG	ACCGGCTCGA	GGGGGTGCGC	GGTGTGGCGG	1140
TGGGTTTCTG	TCGAGTCG	ATCGAATGTC	CGCGCTACCC	CTTCCCCGGA	CTGGGGTGGAGG	1200
CGGGCGCGCG	CGGCGCGCG	ATCGGCTGG	CGATGATGCG	TTTGGGGGAA	CTGCTTGTGCG	1260
CACTGGGTAT	TCCTTACGAC	ACTGAGAAGG	CGTGGGGCTT	AGGCCCCCGG	CTCATCGCTC	1320
CGATACAGCA	GGCGCGCGAC	ACGGGATCGC	GGAGGGCTGG	CGAAGAGCGG	GGCGCGATCT	1380
GGCGCGTCAC	CGTATCGCG	TTTGGGGGTT	CGGGCCCGAG	CGCGCAACGA	CAGGTACCT	1440
CGCTCGCTC	CGACGGCGCA	CGGGGGGTT	CGGGGGGGAG	CGCGCAACGA	CAGGTACCT	1458

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 882 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

ACGGTGTAT	CGTGCTGGAT	CTGGGACCGC	GTGGCCCGGT	ACCTACCGAG	ATCTTACTGGC	60
GGCGCGAGGGG	GTGCGCGCTG	GGCATGGGG	TGTCCTGAGT	CGGGATCGCG	GTGGGCGATCJ	120
TCATCGCTT	CGTCGACAGC	ACGGCGCGTG	CCAAACCGGT	CGGGCGCGAC	AAAGCCGGGT	180
CGGGCGCGAGG	CGTCGCGGGC	TCGGCGCGAC	CCAAACGACC	CGAGCGCGCC	GGGCGAACCGC	240
AAAGCTTAACG	CGCGCGCGCC	CGGGCGCGAG	CGGAAACCGC	CGGAGACACG	ACGGCGACCG	300
CGCGCGCGCG	CGGGCGCGCG	CGGGCGCGAG	CGGAAACCGC	CGGAGACACG	ACGGCGACCG	360
CGCGCGCGAGG	TTTGGCGAC	CGGGCGCGAG	ACTACGCTGG	CGACCGACCG	AGTTTCACCA	420
TGGTGTGTCAC	CGACACGCG	CTGCGTGTCT	GTAAACCGGA	CGTGGGGGGG	CGCGGTGTTGG	480
CGCGCTACGT	TTACTCGCTG	GACACACAGC	GGTTGTGCGC	CAACCTGAC	TGGCGCGCCG	540
CGAAATGAGACG	CGCTGCTGAG	AGGTGTTTCCG	CGGGTGAAGCA	CGTAACGACG	GGCGTGAAGC	600

GGACCGGGAT	GGGATCGGCG	CGGGCTGCC	CATTGCGCGG	GGCGGGGATC	GGGGGGGGCA	660
CCTACAACT	CTGGTGTACAA	CTGGGGCAT	TGGCGTCTGCT	GGGGTTTCCG	TTCATCCTGA	720
ATCAGGGGGC	GGCGGGGGCC	GGGGGGGTAC	CGGCTCGGG	TGAGCGCGAG	GGCGCTCGC	780
CGGAGTCTC	CGCGGAAAGC	GGATAAATTAT	TGATCGCTGA	TGGTGTGATT	CGCGAGCTGT	840
GACAACCCG	CGCTCTGTC	CG				862

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 622 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

TTGATCAGCA	CGGCGGAGC	GTACATGCC	TCCTGGG	TGCGGGTGC	CAATGACAA	50
GACACCCGGG	GGCCGAAAT	CGTGGAAAGT	CTGGCUNSTG	GTGCTGCCUC	GAACCTGGA	120
GTGGCTGAAAG	GGCTCGTTGT	CACCAAGTC	GACCGGGGCC	CGATGACACAG	GGCGGACCGG	180
TTGGTTGCCG	CGCTGCGTC	CAAAAGCC	GGGCCACGG	TGGCGCTAAC	CTTTCAAGAT	240
CCCTGGGGC	GTACGGGAC	ASTHCAAGTC	ACUTGCGKA	AGCGGGAGCA	GTGATGAAAG	300
TGGCGGCGCA	GTGTTCAAAAG	CTCGGATATA	CGTGGGCAAC	CATGGACACG	CTTGCGAGT	360
TGGTGGTGG	CGGGGCGCTT	CTGCTGCTG	TGAGGATCG	CACGGCGCAC	GGCGNATGAAG	420
ACCAACAGGG	GGCGGTGTTG	ACCGAGCTGC	TGACCGGG	CGGGTTTTT	GTGACGGCGG	480
TGGGGGGCGT	GTGCGCCNAC	GAAGTCAGA	TGCGGAAT	GCTGAAACACA	GGGGTGTATG	540
CGGGGGTGG	CGCTGGTGTG	TGCGTCGGG	GGACCGEN	GACGCTTCG	GATGTCACCC	600
CGGAGGCGAC	CGGAGACAT	CT				622

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1200 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

GGCGCAGCGG	TAAGCTT	GGCGCGGCG	ACACTGGTGT	TGACAGCATG	GGCGGGTGGC	50
ACCAACAGCT	CTGGTGTAGG	CGCGAGCGG	ACGTTGGGT	CGGTGCACTG	CGCGCGCGAG	120
AGGGAGCTCC	CTGACAGCGG	CTGAGGCGA	CAAGAAAATG	CGTGTGAGCA	GTGCGTCTAT	180
GGCTTACGGNC	GTATGTCGCG	GGGGTACAGC	TTGAGTACTAC	ACGCCAACGG	GTCCGGTGCC	240
GGGGTGTACCC	AGTTTCTCA	CAAGGAAACCC	GATTGCGG	GTGCGGTATG	CGGGTTGAAT	300
CGCTGGACCC	GTCAACCTGA	CGGGTGGCC	GAOGCGTGC	GTTCGGGG	ATGGGACCTG	360
CGGAGGGTGT	TGCGGGCGG	CGGGTACACG	TACATATCA	AGGGCGTGAG	CGCGCTGAAT	420
CTTGACCGGAC	CCCTACAGCG	CAAGAGTTTC	AAAGGCGACCA	TCACCGTGTG	GAATGATCCA	480
CAATGCGAAC	CCCTCAACTC	CGGCGCGAC	CTGCGCGCA	CACCGATTTAG	CUTTATCTTC	540
CGCGAGGACAA	AGTCGGTGTAC	GTGCGACAC	TGCGAAAT	ACCTCGACGG	TGTATCCAAC	600
GGGGGGTGGG	CGCAAGGCGC	CGGGGAAACG	TGCGCGGG	GGCTGGCGGT	CGGCGCCACG	660
GGGAAACACG	GAAGCTGGCG	CGTACTCGAC	ACGACCGAGC	GGTGGATCAC	CTACAAAGAG	720
TGGTGGTTG	CGCTGGTGTG	CGCTGGTGTG	ATGGGGAC	TGATCGTC	GGCGGGTGGC	780
GATGGCTGGG	CGATGACAC	CGAGCTGGTC	GGTGAAGACAA	TGCGGGGGG	CGAGATCATG	840
GGGAAAGCGCA	ACGAGCTGTG	ATGGACACG	TGCTGTGT	ACAGACCCAC	CGAGCTGGC	900
TCTTACCGGA	TGCTGGTGGC	GGCTTATGAG	ATGGTGTGT	CGAAATACAC	GGATGCGACG	960
ACGGGTTACG	CGGTAAGGGC	GTGATGCGA	GGCGGATTG	GTGCGGGGCA	AGAAAGCTCG	1020

GACCAATACG GCTCCATTCC GTTGCCCCAAAG TCGTTTCAAG CAAATTCGGC GCGCCGGGTG
 AATGCTATTCTTCACTTACG TGAGGGGAT TCAACGCTGA GCGATGGCGT TCCGGAGGTTA
 1080
 99GTGCAAT TTGGGCGCTA TCAAGTAITG CGGCTCTGCG GCGAGGGCGG GATGGCGAG
 1140
 1200

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 1155 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

GCGAGCGAGT GCGAGGCGTG CTGTTTCAACG AATCGGCGAT GCGGAAGACG AAACCGCACCA
 AGACCGGCTA CACACCGAT GCGGACCGCG TCGCTGCGT GTTGGCGCAAG ACCGGGCGATC
 60
 CTGTTTCTGCA ATACGTTT GCGGACCGCG ACCTGACCGG GCTCAAGGTC ACCCTGGCGAC
 120
 GGTGCTCTCA AGCGCTGGCG CGCTGACCGG GCTACGCTCAC GAGACGATCG
 180
 CGCGGACCGG CGCGCTCTCG TGAAGGAAAC CGACCTCTCA GAGACGATCG ATCCGACCG
 240
 ACGGGGCGCG CGGGATGGCG GAGGCGTCA TGGTGGGGGA CGGTTACGCC GAGTTGATGA
 300
 CGGGCGACCTA CGACGAGATC GAGATCGCAAG TGTAGGGCGA CTGTGCGCG GACGAGGGCG
 360
 TCACTGAGGC GTCACACCC GGGGAGGAGC TGTTAGCTGTT CGTCGCGTGC CGGCTGTTGCG
 420
 STGTCGCGGC CGACGAGTC CGGGGGAGGT TGGGCGCCCGG GGTGAGGGCG ATGTTCTTACG
 480
 GCGCTGTTTA CGGGTTGAGC CGCTACGCGT TGTCGCAAGCA GTTGAATTC TCAACGCGAG
 540
 AAGCCAAACCGA CGACATGAGC GCTTATTTGC CGGATCTCG CGGGGTGGCG GACTACCTGC
 600
 GCGCCGTAGT CGACGGGGCG CGCAAGGAGC GCTACGCTTC GACGCGTGC CGCCGCTCGG
 660
 GCTCTTCCCG CGACGAGTC GAGACGACCC GTCAGTGGCG GGAAGCCCCCG ATGATCCAGG
 720
 CGCTGAAACCG GCGGAGTCAG GCGACGGGGG CGACATCAT CGAAGTGGCG GAGGCGGGCG
 780
 TCGACAAGGC GTCACACGC CGACAGCTGG CGTCGCGAT GCTGCTCGAG GTTCAAGAGC
 840
 AGCTGCTGTT CGAATTCGGCG CGCGGTGAGC GCGAGGGCGT CGAGGGCTCG GGGCGCGAG
 900
 AGATGGCGCG CGCTTACCGC CGCTACCTTC CGCTGAGGT CTGGTGGCG TACGGCGCGA
 960
 CGCTGGGACCG CGCGCGCGAC TGACTGGCGG GCGTGCAGCT CGGGCGGCGA TTGGCGCGATT
 1020
 TTGGCGCGTT CGGGCGCGAC TCGGGCGAT CGGGCGCGAG TTGGTGGCGC GTTACCGCGT
 1080
 CGAGTAGCGT CGTCGA
 1140
 1155

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 1771 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

GAGCGCGCGT TGGTGTGTTGA AGCGTTTTCG CGGTGCGCAT CGGCACGGCG GTTGGCGGGT
 TCGGGCGCGTGG GTTGGCGCGAT CGTCAAACAGG GTGGTGTCTCA ACCACGGCGG ATGGCTGCGC
 60
 ATGGAAAGACA CGGCAACAGG CGGCAACGGCG CGTGGACGCT CGATTCAGCT GCTGCTCGCC
 120
 CGGCGTGGGA TGGCGATTCC CGACGTTCC CGTGGCGCG CGGGCGCTCG GAGCACGGAC
 180
 ATCGAAGACT CGCGGGGTT CGCGACCGT ATCTCGAGTG ATCTCGAGTC CACGGCGCGA
 240
 ACCTTATTTG CGGTTACGTC TTGAAGGCA CGACCGTGGC AGTCACCGCA TGGCCAGTT
 300
 CGGCGAGTA CGGGGGTTAG TACAGGAGA GCAACTTACG GACATGACCA ATCACCGCG
 360
 GATTCGCGA CGCGCGAGC AGCCGGGAAC CGCGGGTT GCTGAGGGCG AGCACGAAAC
 420
 CGACGCGCG AGTTCGACT CGGCTTACCG ACCTTACCGG CGCGCGCGC CAACCGAGTA
 480
 CGCTGAAACCG TACGGACGCT TGGTGTGTTAC CGGGCGGGT CTGATACCGT CGCTGATTC
 540
 600

GACCATGAGC	CCCCCTCTG	GGATGGTTG	CACAGGCGCT	CCTGCAGGCA	TGTTGGCAT	560
CGGGCGGTTG	ACGATAGCGG	GGGTGTCGCG	CACCATGGC	GGCGCGGGCG	CATCCCTGGT	726
CGGGTTCAAC	CGGGCACCCG	CGCGGCGCG	CGGGCGGCGCA	GTGGCTGCGCA	CGCGCGGGCC	780
AACCATCCCC	CGACCAAAAC	TGGCGCGCGG	GTGGCTGCGAA	CAGGTGGCGG	CCAGGTGGT	840
GGCCATCTAC	GTCACTGTTGG	AAACCGATCT	GGGCGGCCAG	TGGAGGAGG	GTCCCGGGCAT	900
CATTCTGCT	GGCGAGGGGC	TGAGCTTGAC	CAACAAACCC	GTGATGCGCG	GGCGCGGCCAA	960
GCCTCCCTG	GGCAGTCCCG	GGCGGAAAAC	GAACGGTACCG	TTCCTCTGACG	GGCGGACCGC	1020
ACCCCTTCAGG	GTGGTGGGGG	CGTACCCCGAC	CACTGTATAC	GGCGTGTGTC	GTGTTTCAAGG	1080
CGGTTCTGGG	CTTACCTCTG	TTCCTCTGGG	TTCCTCTGCG	GTGACCTGAGG	TGGGTCTAGCC	1140
GGTGTCTGGG	ATCGGGTGC	GGCTGGTTT	GGAGGGCACC	GTGACCCACGG	GGATCTTCAG	1200
CGCTCTCAAC	GTCTCACTGT	CGAGGACCGG	CGAGGCGGCC	AACAGAAAC	CGTGGCTGGA	1260
GGCCATTCTAC	ACCGAGGCCG	CGATCAACCG	GGCTAACTCC	GGGGGCGCGC	TGGTGAACAT	1320
GAACGCTCAA	CTCTCTGGAG	TCAACTCTGGC	CATTGCGAAC	CTGGCGCGCG	ACTCAAGCCG	1380
TGCGGAGAGC	GGCTGATGATG	GTCTGGGTTT	TGGATTTCA	GTGCGGACCGG	CCAAAGCGAT	1440
CGGCGACGAG	TGATCTAGCA	CGGCGAACGC	GTACACAGCC	TGCTGGGTG	TGAGGGTAC	1500
CAATGACAA	GACACCCCGG	GGCGGAAAGAT	GTGGCGAAGT	GTGGCGCGTG	GTGCTCCCGC	1560
GAACCGTGA	GTGCGCAAGG	GGCTGGTGT	CACCAAGGTC	GACAGGCGCC	CGATACACAG	1620
CGCGGACCGG	TTGGTGGCGG	CGTGGCGCG	AAAGGCGCCG	GGCGGCGCGG	TGGCGCTAAC	1680
CTTTCAGGAT	CGTCCCGGGG	GTAGCGCGCG	AGTGCAGTC	ACCTCGGCA	AGGCGGAGCA	1740
GTGATGAGG	CGGCGCGCGA	GTGTTCAAG	C			1771

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1656 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) SEQUENCE DESCRIPTION: SEQ ID NO:14:

CTCCACCGCG	GTGGCGGGCG	CTCTTAACT	AGTGGATGCC	CGGGGCTGCA	GGAAATTGGG	60
ACGAGGATTC	GAACCTGCG	GTGGTGGACG	CGGGGCGCG	GGAAAGTATCG	GTCCCATGCG	120
AGCCCGGGCG	GGGCGACCGC	CGGAAATGGCG	CGAGTGGAGA	GGCGGCGCA	TTGGGGGGCG	180
CGCGGACGCG	CGAGGCGCCG	AACTGGCGCA	GTGAGGAGGC	GGGCACTCAT	GGCCAGCGTG	240
ATCCCATCAA	CTTCGACCG	GGTGGGGCG	CGATTGCGCA	ATCGAGGTAG	TGAGCGCRAA	300
TGATATGATG	AAACAGGCGG	GTGAGGCGCG	CGATTGCGCA	GTGAGGAGGT	GGCTGGCTGG	360
CGTTTGGGT	ATCAGGATGT	TCTTTCGCGA	CGATTGCGCA	GTGAGGAGGT	GGCTGGCTGG	420
TGAGCGGAC	GGGCTCCGAC	CGGGGCGCG	CGAGGAGACG	GGTTTTCGG	GGGCTCCGAC	480
CGAAAGGGTT	GACCGCGCTG	CGCTGAGCGG	TGGCGGAGAT	CAGGCGTCG	GTGATGCGA	540
TGGGTATTAC	CACTGCGGAT	GTGAGCTGGC	CGGGAAAGTC	CGGGAAAGTC	GACAGCTTCG	600
GCACCTTACAA	CGACGCGAC	GGTGGCGGT	GTCTCCGGCA	AGGGGCGTAT	ATCTCGGTGA	660
AACTGTTGCA	CGACTGGGAGC	AATCTGGGT	TTGGGGTACA	AGGGGACAAAC	ATCTCGGTGA	720
TGGATCTGC	CGCTGGGGGT	ACGAGCTG	TGGTGGGTGT	ACGAGCTC	TCACGGCTGC	780
TCACCGGAGT	GTAGCGCGA	ATTCGACCA	CGAAATTCAC	CGGGGACCATC	CGGGCGAGCT	840
CTGTCAGGAT	GCTTGATCT	GGGCGGAGCA	GTGCAAGGCC	GGCGACCGTG	TGGATGCGCC	900
AGGACGGCTC	GCACCGACCT	GTGCGAGCGA	GTGATGCGGT	CGGATGCGGG	TGAGTTGCGC	960
TCACCGAGTC	GAATGCGAC	GAACCGCTCA	ACGTCGACTA	GGCGGAGTT	GGCTCGACCC	1020
GTGTTGCGAA	ACCCGCTCTG	GAAGGGTGT	AACGGNAC			1088

(2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 542 base pairs
- (B) TYPE: nucleic acid

(C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

GAATTCCGGCA	CGAGAGGTGA	TGACATCAT	CGGACCCAGC	CCCACATCTT	GGGAACAGGC	69
GGCGGGCGAG	GGGGTCGAGC	GGGGGGGGGA	TAGGTCGAT	GACATCGCGU	TGGCTCGGGT	120
CATTGAGCGA	GACATGCGCG	TGGACAGGCC	CGCGNAGATC	ACCTACCGCA	TCAAGCTCGA	180
AGTGTGCTTC	AAAGATGAGGC	CGCGTACAC	GGCCCGGGCA	GCAAGACGCA	240	
AAATCSCACG	GTTCGCGTT	ATTTTGTCG	TGCTCGCGA	GGCTTACCG	300	
GC CGCGCGCG	GTGCGCGCGT	CTCGCGTATC	CGCGCGTGC	GGCGCGGCG	360	
CGCGAGTTAA	TCTTCGCGT	CGACCGGAC	TGGCGATTC	TGATCGATGA	420	
CGTGGCCAG	CGCGTCCGAT	CTCGAGTTGC	CGAGGAAAC	GTGCTCCCG	480	
AGCGCTTCTA	GGCGCGCGTG	CTGACCGCGT	CTCTAGTGGG	GGCGAGGAGC	540	
GG						542

(2) INFORMATION FOR SEQ ID NO:16:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 913 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

GGGTGCGCGC	CGGGCTCCG	TTGGCGCCAT	TGCGCGCGTC	GGCGATCAGC	TGCGCGATCG	60
CACCATCACC	GGTTTTCGG	CGGGCACCGC	GGTGTGCGCC	GGGGCGCGCG	ATGGCCACCGC	120
TTGACCTGG	CGCGCGCGCG	GGCATTCGCG	ATACAGACCC	GGGGCGCGCG	CACCTTACG	180
GCCTGCGCGC	CGCTCGCGCG	GGCTGGCGT	TCACCGCGCG	GGGGCGGGGG	GRACCGCGCG	240
CAAGCGCGCG	GCACGACCGC	TTGCGCGCGT	TTGCGCGCGT	GGGGCGGGCG	CGCGCAATTG	300
CGGACACCGC	AMGACCGC	GGCGCGCGCG	GGCGCGCGCG	TAACGGGGGT	GGGGGGGGCG	360
GGCGCGCGAC	CGCGATTCG	CGCGCGCGCG	GGCGCGCGCG	GGGGGGGGCG	GGGGGGGGCG	420
GTTCGCGCGC	AAATATTCGG	GGGCAAGCGC	AGACCGCGCG	GGGGCGTTAC	GGGGCGCGCG	480
CACCGGAAAC	ACAGCGCGAC	GGTGCAGCGC	GGCGCGCGCG	GGGCCCCAT	TGCCCCCGGG	540
TCACTGGCGAG	CACCGCGCGT	AAATTTTATG	ACCGCGCGAC	TTGCGCGCGT	CACCGGGCGAT	600
CGGGGCGCGC	GRAGCGCGTC	CGCGCGCGCG	CGCCAGCGCC	GGGGCGGGCG	GGGGCGGGCG	660
CGGGCGCGCG	GGACCGACCG	GTGCGCGCGA	TCACCGCGCG	GGGGCGGGCG	GGGGGGGGCG	720
TGTTGCGCGT	GGACCGACCG	GGGGCGCGCG	GGGGCGGGCG	GGGGCGGGCG	GGGGGGGGCG	780
CGGGCGCGCG	GTGCGCGCGT	ACCGACCGCG	GGGGCGGGCG	GGGGCGGGCG	GGGGGGGGCG	840
TGGCGCGCGT	GGGGCGCGCG	GGGGCGGGCG	GGGGCGGGCG	GGGGCGGGCG	GGGGGGGGCG	900
CGGGGGCGCG	CGCGCGCGCG					913

(2) INFORMATION FOR SEQ ID NO:17:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 1872 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

GAATACGTTG	GTGTTAGAAAA	ATCTGCGCG	CGGACCGCTT	AGGGCTGGGA	CAATTTCTTA	60
TAGCTTACCG	GACACAGGAG	GTGAGGGAT	GGGGATTCG	GGGGCGCGCG	GGGGCGGGCG	120

GTCACTGGTTG	CTGAGCGTGC	TGGCTGGCGT	CGGGCTGGC	CTGGGCCACGG	CGCGGGCCCA	180
GGCGGGCCCGG	CGGGCTGGTG	CGCAGGACCC	GTTOGGCGAC	TTCCCCGCGC	TGCGCGCTCGA	240
CCCGCTCGCG	ATGCGCGGCC	AAAGTGGCCG	ACAGCTGGTC	AACTCTACCA	CCAACTCTGG	300
CTACACACAC	CGCGTGGCGT	CGGGGACCGG	CATCGTCATC	GATTCACAGG	GTGTCCTGCT	360
GACCAACACAC	CACCTGATCG	CGGGCGCCAC	CGNCATCAAT	CGTTTCAGCG	TGCGCTCCCG	420
CCAAACCTAC	GGCGCTCGATG	TGGCTGGGTA	TGACCGCCAC	CAGGATGTCG	CGGTTCTGCA	480
GCTGCGCGGT	GGCGGTGGCC	TGGCTCGGG	GGCGATCGGT	GGCGGGCG	CGGTTGGTGA	540
GGCGCTCTCG	ACASCGTGG	GGCGGGCGA	ACGGCCCGTC	CGGTTCTCGG	600	
CAGGCTGTCG	GGCGCTGGCC	AAACGCTGCG	GGCGTGGGAT	TCGCTGACCG	GTGCGCGAAGA	660
GACATTGAAAC	GGTTTGTATTC	AGTTGGATGC	CGCGATTCAGG	CCCGCTGATTC	CGGGCGGGCC	720
CGTCGTCACAC	CGCGCTGGAC	AGGTTGGTGG	TATGGACAGC	GGCGCGCTCG	ATAACTCTCA	780
GCTGTCCTCG	GTGGGGCAGG	GATTCGCGAT	TGGATCGGG	CAGGATGATCG	CGATCGCGG	840
CCAAATCGGA	TGGCTGGGG	GGTCACCCAC	CGTTCTATTC	GGCGCTACCG	CTTCTCTCGG	900
CTTGGGGTGTG	CTGGACAAACA	ACGCGCAACG	CGCACCGAGTC	CAACCGCTGG	GGGAGGCGC	960
TCCCGCGCGCA	ACCTCGCGGA	TTCGACCGG	CGACCTGATC	ACCGCGCTCG	ACGGGGCTCC	1020
GATCAACTCG	GGCGACCGGA	TGGCGACGCG	CGTAAACGGG	CGATCTCCCG	GTGACGCTCAT	1080
CTTGGGGTAAAC	TCGCAACACCA	AGTGGCGCGG	CGCGCGTACA	GGGAACTCTA	CATTCGGCGA	1140
GGGASCGCGG	ACCTGGATTG	TGGCGATCG	CGGGCGCGG	CGCGCGATTG	GGATTGGCGC	1200
CAGCGCTGAT	TGGCGCTGGA	GGCGCGCTGT	TGGCGCTCGG	GTGGCGCTGG	CATTGTGGAA	1260
GCATGACACG	AGGCGACACA	CGCGGTGAG	CGACCTCTGG	TGCGGGCG	TTACGCTGAA	1320
GGCGGTGG	TGCGACATCG	GGATTCGCG	CGTACCGTGG	GGCGCGCGC	CTTGGCGGCC	1380
GAATCGGACCT	GGTTTAAAGCA	CGCGCTCTTC	TAGGAGGTCG	TGTTCCGGGC	GTTTTTCGAC	1440
GGCGCGCGG	ACGGTTGGCG	CGATCTCGGT	GGACTCATCG	ATCGCTCTGA	CTAACCTCGAC	1500
TGGCTTGGGA	TGGACTGCG	CTGGTGGCGC	CGTCTCTCG	ACTCACCGT	GGCGCGACCG	1560
GGTTTAAAGCA	TTCGGGACTT	CTACAGGTC	CTGGCGGAAAT	TGGGACCGT	CGACGATTTC	1620
CTGCGCGCTGG	TGCGACACCG	TGCGCGCGA	GGTATTCGCA	TCATACCGGA	CTTGTGATG	1680
ATTCGACACT	GGGAGGCGA	CTCTGGTGT	CGGAGGTCG	GGCGCGACCG	AGACGGACCG	1740
TACGGTGTACT	ATTACGTTG	GGCGACACCG	AGCGAGGCT	ACACGACCG	CGGGGATCGT	1800
TTGGTGGACA	CGCGAAGAGTC	GAATGGTCA	TTGGATCTTG	GGCGCGACCA	GTTCGCTACTG	1860
GCACCGATTC	TT					1872

(2) INFORMATION FOR SEQ ID NO:18:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1482 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) SEQUENCE DESCRIPTION: SEQ ID NO:18:

CTTCGCGCAA	ACCTGATGCC	GGGAAACAGG	GTGTTCCCGT	GGGGCGGAGG	GGGTGCGGAC	60
CCCGCGCTCT	CGCGAGAGTC	AGCGAGTCGC	TTGATGCGAC	AAAAGGTTTG	ACCAAGGCGC	120
ACGTTAGCGGT	CGGAAACACG	GGGAAAGCTG	ACASCTCTCT	GGGTATTACG	ACTGCGGATG	180
TGACGCTTGC	GGCGATCGAC	GGGGCGCTCG	CGCTTACAAAC	GAACGAGCAGG	GTGCGCTGGT	240
GTGTCGCGTT	TGGGGTACAA	GGCGACACAA	TTCGGTGTGA	ACTGTTGCG	GACTGAGACCA	300
ATTCGCGCTC	GATTCTCTGA	CTGTCGACTT	CGCGCGTCT	CGATCTCGCC	GCTGGGGGTGA	360
CGCAGCTTGT	TGCGCGTGT	ACGAACTTCC	AAACGCAAGG	TACCGAAGTG	ATAGACGAAA	420
TTTGCGACAC	CAAAATCACC	GGGAGCGTAC	CGCGCGAGTC	TGTCGAAGTG	CTTGATCTCG	480
GGCGCAAGAG	TGCAAGCGG	CGCGACCTGT	GGATTCGCGA	GGGACGCTCG	CAACGCTCTG	540
TGGCGACGAG	CATCGACCG	GGATCCCGGGT	CGATTCGAGT	CAACGAGTCG	AAATGAGACG	600
AACCGCTCAA	CGTCTCTAGT	GGCGAAGTGT	CGTGGACCGG	TTGGCTGAAA	CGCCCGTGTG	660
AACCGTGTCA	ACGCGCGCCCG	AAACCTGACG	CGTGGACCGG	ATCTGAAAT	TGACCCCGTA	720
GACCGGGCGG	TTGGCTGTTA	TCTCTGGTG	GTGGCGCTG	GTGGGAGCGG	GGCGAGGCTG	780
CGCTTCTTGA	GGCGGTAGCT	GTGCGCTTGTG	AGGGGGCGCGA	CTTCACCATG	GTGGACGAGG	840

CCTGGTGGATCA	TGGCGGGCGAC	AACGGACGTCG	TCCCGCGCGA	AAACCTCGCC	CCACCGGGCG	900
AAGGGCTTAT	TGGACGTGAC	GATCAGCTG	GGCGCGCTAT	ACCGGGAGGA	CACCAAGCTG	950
AAGAAAGAGT	TGGCGGGCGTC	GGCTCTGAC	GGATGTAAC	COACTTCTC	AAACCAACAGG	1020
AGCGGATAGC	GGCCAAACCG	GGTGAAGTCG	GGCTAGATGC	GGCGCGCTG	GTGAAGCTCG	1080
GGCGAACCTG	CTACCCATTG	GGCGCGCTG	GGCAACAGCA	GGCGATGAC	GGCGCTGACAC	1140
GGCGCTATCG	CGAGGGCGAC	GGCAAGATG	GTGTTTCCCG	TGCGAGGGCG	GGCCCAAAAAA	1200
CACGACGTTG	TGGCGGGCGG	TTATGAAATC	CGGGTGGCC	AGATGTGCGA	TGGTGTGCG	1260
TTTGAGGGCA	CGACGATGCT	CAAACTGAA	CTCTTCCAC	GACTTCCGAA	CGCGGAGGG	1320
GGCGGGCGGG	ATNSGGCGCT	CCACCAATG	GGACTCCCG	GTGACACTT	CCCGCTGCG	1380
CGAGGGCGCC	AGGTATTCCT	CTTGCGCTCA	GTTCGCGGC	CGGCGCGAT	CGGCCACCG	1440
GGACACGTAC	TCACCGAGGG	TGGGAGCTTT	CAATGCTCTT	GT		1482

(2) INFORMATION FOR SEQ ID NO:19:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 876 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

GAATTCGA	CGAGCGGGCG	ATAGCTTCTG	GGCGCGCGCC	GACCAAGATGG	CTGAGGGGTT	60
CTTGTGCGG	GGCAACCGCG	GGCGCACCAAC	GGTACCGCGT	GGGGGCTTC	AACACCGCGA	120
CGGTCACTCG	TGTCGTGCTG	ACGGCACCAA	GGCGCGCGT	TTTGGCTTACG	ACCGGGCGTT	180
CGCGTCTGAA	ATTCGCTTACA	TTGCGAAGG	GGGACTCGCC	AGGATGTGG	GGGAGAAACC	240
GGAGAACATA	TTTCTTCTAA	TCACCGTCTA	CAACGGACCC	TACGTGCGC	CGCGGGAGCG	300
GGAGAAGCTC	GTACCCGAGG	GGCTGCTGGG	GGGATCTTAC	CTTATACCG	GGGCCACCGA	360
GGCAACGACCC	AACAAAGGNG	AGATGCTGG	CTCCGGGGTA	GGGATGCGG	CGGGGCTGCG	420
GGCGACGACG	ATGCTGCGG	CGCAATGGGA	TTGCGCGCC	GGCTGTGTGT	CGGTGACCG	480
TTGGGGCGAG	CTTAAACCCCG	ACGGGGTGTG	CTACGAGACG	GAGAGCTTGT	GGCACCCCGA	540
CGCGGCGGCG	GGCGTGTGCT	AGCTGACAG	AGCGCTGGAG	ATGCTGCGG	GGCGGTTGAT	600
CGCGGCGGCG	GTCTGGATCG	GGCGGCGT	GGAGCAGAGT	GGACCGTGGG	TGCCGGCGAC	660
ATACCTTCAGC	TGGCGACCG	ACGGGGTGGG	TTTGTGCGAC	ACTGGCGCCG	CGCGGTGCG	720
TTTACTTCAC	ACCGACCGCG	AAATCCAGGT	TTGCGCGGT	TTTGGGGAGGG	GTGCGCGCG	780
TTGGACGGGTG	AATATGCGAC	CATTGGGGTC	CGCTGCGGGG	CGGCCCGCGCC	AGTTACCGGG	840
ATTCGACGAA	GGTGGGGGGT	TGGCGCGCGA	TTAGTT			876

(2) INFORMATION FOR SEQ ID NO:20:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1021 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

ATCCCGGGG	GCTGCGAGGA	TTCGCGACGA	GAGACAAAT	TTCACCGCGTT	AATGCGAGGA	60
CAGATTCTA	ACGAATTCTA	ACCGGGACAA	CAATATGTCG	GGATGCGGT	TTATTCGAC	120
ACGGAGAGC	TGGCGCACTT	GGCGGAGCT	TTTACGAGCC	AAAGCGCTGA	GGAAAGGAAAC	180
CATGGCGATG	TGCTGTGCTA	ACACTGCTC	GGCGCGACG	TTGGTGTGCGA	AATTCGGCGC	240
GTAGACGACG	TGGCGACGAA	GGTGGACAGA	GGCGCGGAGG	CACTGCGCT	GGGCGCTGCG	300
CGGGAAACGA	CAGTCACCGA	GGAGTGCGT	CGCTGACAG	CGTGTGCGG	CGAGGAGGGC	360
GATTTCCTG	GGAGGCGATT	CATGCGATGG	TTWTTGCGGG	AACAGATCGA	AGAGGTGCGC	420

TIGATGCCAA	CCCTGGTGCG	GGTTCGCCAT	CGGGGCGGGG	CCAACCTGTT	CGAGCTAGAG	480
AACTTGTG	CACGTGAAGT	GGATGGCGC	CGGGCCCGAT	CAGGGCGCCC	GCACGCTGCC	540
GGGGGCCGCG	TCTAGATCCC	TGGGGGGGAT	CAGGGAGTGG	TCCWTTTGC	CGCGCGCTCC	600
TCCAGCCAGG	CCTTGGTGC	GCCTGGGTTGC	TGAGTACCAA	TGAGGGCCAC	CGGGAGCTCC	660
CGGNNAAAGT	CGATGTCCTC	GTACTCATCC	ACGTTCCAGG	AGTACACCCG	CCGGGCGCTGA	720
OCTGGCCGAGC	GTCAACGAG	TTCGGGATAT	TCTCTTACCG	CAGGCAGTGA	GGGTGCGGACG	780
GGGGTGGGCC	CGACGCCGCT	GGCGCGACTG	CTGTCAGGT	ATCGGGGGGT	CTTGGCGAGC	840
AACAACCTG	CGAGGAGGGG	TGAGGCGGC	CGATGCGCA	GGGGGGGGG	CGAAAACGCA	900
CATCAACACG	CGACGGGATC	GATCTGGGA	GGGGGGTGC	GGAAATACCGA	ACCGGTTGAG	960
GGAGGCCAGC	AGTGTGTTT	CCACCGAGGA	AGCGTTTTCG	GGTCATGGN	GGCINNTTAAG	1020
T						1021

(2) INFORMATION FOR SEQ ID NO:21:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 321 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) SEQUENCE DESCRIPTION: SEQ ID NO:21:

CCTGCCGAGC	AACGGAAAGA	CACAAACCATG	AAGATGGTGA	AATCGATCGC	CCGAGCTCTG	60
ACCGGCCCGG	CTTCAATCGG	CCCGCTGCG	GGCGGGTGTGA	CTTCGATCAT	GGCTGGCGGN	120
CGGGTGTGAT	ACCRGATGCA	GGCGGCTCCTC	TTGGCGGCGG	CACTGCGCTT	GGACCCGNGA	180
TCCGGGCTTS	ANGTCGGAC	CCGCCCGGAG	TGGACCGAGC	TGCTAACAG	NGCTGCGCAT	240
CCCAACGCTG	CTTGTGNGGA	CAAGGNGATG	CTGGTGGAGG	GGNGNATCGG	GGNGNAGGAG	300
GGNGNGNATC	GNCGANACAA	A				321

(2) INFORMATION FOR SEQ ID NO:22:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 173 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) SEQUENCE DESCRIPTION: SEQ ID NO:22:

TCTTATCGGT	TCGCGTTGG	GGGGGGGTTT	GGGGGGGGGT	GGTTAACCCG	CTGGCCGAGC	60
CGATGACGG	GGGGGGGAGAC	GTCACTCGG	ATACTCGCG	GGGGGGGGAG	CTCCAGCGCC	120
CCTCGGGTGT	GNACCGGCAA	GGCGGGAAAG	ACCGGTTGNA	GGGGGGGAGC	AAGGCGATTG	180
ACCGGATGAC	CCCGATCGGG	CGGGGGCAGC	GGCACTGTAT	CATGGGGGAC	CCGAGACCG	240
GCAAAACCGC	CGGTCTGTGT	GGGACACCAT	CCTCAACAC	GGGGGAGAAA	CTGGGAGATCC	300
GGTGGATTCG	AGAGGAGGAG	TGGCTTGATG	TATACTGGTG	CCATGGGGCA	AGAAGGGGAA	360
CTTACCATCG	CGG					373

(2) INFORMATION FOR SEQ ID NO:23:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 352 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

GTGACGCCGT GATGGGATTG CTEGGCGGGG CGCGCCGCT GGCGGTGGTG GATCACCAAC	60
TGGTTACTCG GGTGCGCMA GCGTGGCTGT TTGGTCGAGC AGCGCGTGTG CGCGTGGTGT	120
TCTTGCAGGC CTGGTACGGG TTGGCGGATT TAGCGAGAT CGAGCGGGGC GAAATGGTGC	180
TGATCTACG CGGTACCGGC GGTGCGGCA TTGGCGCTGT GCAGCTGGT CGCGAGTGGG	240
GGGTGGAGGT TTGGTCAAC CGCACGCCG GNAATGGGA CACGCTGCAC CGCAGTGGT	300
TTGACAGCA NCCATATGG NGATTCCTCC ACATCGAAG TTCCGANGGA GA	352

(2) INFORMATION FOR SEQ ID NO:24:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 726 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

GGAAACCCCG TTCACTCCGT TCGACCCAGG CTGCGGATA ATCGACGAAG TGATCAAGCC	60
CGGGTTGGCG CGCTCTCATG GTCACACCGA GAACTCAGCA AGTTCTCTGG TATATCGAC	120
CTAACGTCGCA GTTGTCTGCA AGATGGCTTT CGTACCGCTCA TCGCATGTCAC CGTTTGGCGT	180
CGCCGACGCT CTCGCTGGCG CGCTGCGATCC TGGCGACGGG TTGGCGCGGT CTGGGGGTG	240
GGCGCCGATC CGCAGCCCAA ATCGCCCGCG TCGCGACTA CTACTGGTGC CGGGGGCAGC	300
CTTGGCGGGG CGCATGGGGG CGCAACTGGG ATCGCTACAG CTGCGCATGAC GACTTCGACC	360
GGCACACCGA CGGCCCGGAC CACAGCCGGG ACTACCCCGG ACCCGATGCG GAAGGTTCCCG	420
TGCTTGACGA TCGCGGTGCT CGCGCCCGCG CGCGCGCTGC CGTGGCGGGG GCGTACGGCT	480
CGTGGCGCG CGCGCGATCG CGTAACTGGG TATAAAACCGG CGCGTGGCG CGGGCAAGCTA	540
CGAGCCCGGG CGGGGCGAGAT TTAGGCTCG CGTGGCGATGG ATCGCGCGGT CGGATGACAG	600
AAAATAAGCG AGCGTTTGG CGACCGTGG GAGGAGCGTT GAAAGGACCC TGTCTGAAAC	660
GGCGACAGCG CGTGGCGACAT CGACATGACG AAGGTTGTTA CGCGCACACCC CGTGGCGCGG	720
ATCGTG	726

(2) INFORMATION FOR SEQ ID NO:25:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 580 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

GGCGACGAGC AGCAACGTCG GCGTCACCGC CGCGTATGGG TTGATGAGG CGACCGGGAT	60
GGTCGCCGAC CATACTCAAG CATGCTGGGT GCGCGCTGAG CGACCTTGTG ACCAGCGGGG	120
CTGGCGATG CGCGCGCGGT GAACTCTATG CGCGCGGGGT TTGTCACCTG ATGAACTCGA	180
ATAGGGAACA ATAGGGGGGT GATTTGGCG TTGATGTTGGG GGTATGGTGT GAAATCCAT	240
GGCGGGGGAT GCTGGCGCGD GACCGAGGCTC CGCGACGGCG CGCGACGGCG ATTCGGAGGG	300
AGCACTCAAT GGCGCGATCG AAGCCCGGGA CGCGCACACG TCGTTGSSAA GCACTTAAGG	360
AGGGCGGGG CATTGTGATG CGATGACACG TTGGAGGTGG CGCGCGCTG CGTGGCGAGC	420
TGACACCGGA CGAGCGGGG CGACTGGGTG CGCAACTCAA AGGCGTTACT AGCTAAGAC	480
AGCCCAACGG CGATGGTGTG CGTAACTGGG CGACACTCAA AGGCGTTACT AGCTAAGAC	540
TGCGCGATGT ATGCCAGGA GAACTTGGG ATACAGCGCT	580

(2) INFORMATION FOR SEQ ID NO:26:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 160 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

AACCGAGGCG CGCGGGGTTT	TGGGGGGGCG	GGCGGGGTG	GGGCAACCG	CGGGCGCGC	60
GATACCGCGG	GGTTGGTCGG	TTCGGGGG	GGCGGGGGG	GGGGAGGCGA	120
GGTGTACCGG	GTACGTGGC	CAACACCCG	GGTGGATCCG		160

(2) INFORMATION FOR SEQ ID NO:27:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 271 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

GACACGGATA CGATGGTGT	GTACGGCAC	GGTTCGACA	GGCTCGAGG	GGTCAGATC	60
CACCGCACAC	CGGAGGGGT	GACCATGGC	GATCGGGCC	GGCGGGCTGCC	120
AAGGCGATGG	GAATGGACAA	GCTCGGGGT	CGTTGCGGA	GGGGGGCTGCC	180
GAACGCGAAC	AGTGGGACGA	CGGCAACAC	ACGTTGGGT	TGGGGGGCGG	240
GGCTACGGAC	GCAACGTACA	GGACACCGC	TGTGGTTTC		271

(2) INFORMATION FOR SEQ ID NO:28:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 317 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:28:

GGAGCGGTG	GTTCCTGGAC	TATTCGGCA	GGTGTACCGA	GGCGGACGTG	CGCGGAGCTGA	60
AGCGGATCGA	GGACGACGT	CGCTGGCGG	GGTTCATCGG	CTACCTGGCC	GGTATCACCC	120
CGCAGGAGCT	GAACGTGGCC	GAAGGGCGC	GGGTCACTGG	GGTGGACGCC	GGGACGATCC	180
GTTCGGATCT	GGCGTGGTTG	GAGACGGTCT	ATCTGGTACA	TGGCTCGCC	GGCTGGCTCC	240
GGAAATCTGAC	CGCGAAGATC	AGAAAGCGT	CTAACGATCCA	GGTGGTCAC	AGTGGCTTCG	300
CGGCGCTGTT	GGCGGGG					317

(2) INFORMATION FOR SEQ ID NO:29:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 182 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:

GATCGTGGAG CTGTGATGAA ACAGCGTTGC CGGACGCCU GCGGCCAGCA CTTGGTATA
 CGACGGCGGG ACCACCTCAG CGTGGGCAAG CAGCTGATG ACCACGTTGG CCTGGGCCAC
 CGCTTCGGCGG CGCTCTACGAA ACACCGGAC ACCCTGGCGG GCGCGCGGG AGCGCGCGT
 CGG 60
 120
 180
 240
 300
 360
 420

(2) INFORMATION FOR SEQ ID NO:30:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 308 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:30:

GATCGCGAAG TTGGTGAGC AGCTGGTGA CGCGAAAGTC TGGCGCCCTG CGAAGCGGGT
 CGCGCTTCAC GAGGCGAAGA CACCGCTGTC CGACCTGCTG CGCTCTGCTT ACAGGGGGCR
 GAGGTTGAGA TTGGCGCGCGG CGGGGAGCGG GTAGCAANGC TTGTGCGCGT GCATCTCTAT
 GAGACTGGC GTTCTGGCAT TGTACATGGC GTTACCGCG TCCCGGACGA TTGTACGCGT
 CGCTGTGAG AGCAGCTCTG CGACGCTTTC CACCGCTGAA GCGCTACCTC ATCGACACCC
 ACCTTGG 50
 120
 180
 240
 300
 360

(2) INFORMATION FOR SEQ ID NO:31:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 267 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:31:

CGGACGACCA CGAACTCAGG TGGATGATGG TCGGAGCGG CATTGAGGAC GGAGAGAGTC
 CGGCCGAAAG CGCGCGCGCG CAGCTGCTCA TACTGACCGG CGGTAGAGGG CTCCCCCGAT
 CGCACCGGAC TATTCTGGCT TCGCTGGC CGTAAAGAGC CGGTAAGAAAGA ATGTGAGGGG
 ACACSATGAC CGATCACACC TACCGAGTC TEGAGATGT CGGGACCTCG CGCGACCGCG
 TCGACCGCGG AATCCAGGGC GCGCTGG 50
 120
 180
 240
 300

(2) INFORMATION FOR SEQ ID NO:32:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1519 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:32:

CTGCTTCCGA AAGAAATGTA GGGGACACGA TGACGAACTCA CACCTACCGA GTGATGAGA
 TGTGGGAC CTGCGCCGAC CGCGCTGAGC CGGCAATCA CGGGGGTCTG GCGCGAGCTG
 CGCACGACCT CGCGCGCGCTG GACTGTGTCG TACTGAGTCG AATTCGAGGC CACCTGGTCG
 AGCGAGCGGT CGCGCGCTTC CAGGTGCTCA TGAAAGTCG CTTCGGCTGG AGGATTCCTG
 AACCTTCAAG CGCGCGCGAT AACTGAGGTG CATCAATTAG CGACTTTTCC AGAACACCT
 GAGCGCGCTG AACCGCGGTT CAGCTGACGG TGGCTGGCG GAGGGCGCTCG CTCCRAAATC
 CCTGGACCAA TTGTGCGCGG CGCGCTAACAA GGAAGTCGTG GCTGAATTG TGCGGTAATC
 60
 120
 180
 240
 300
 360
 420

GGTCGACCTG	TGTGGGGCTGC	AGCCGGAGCA	ACCGGCTGCTC	GACGTGGGCT	GGCGCTGGG	480
GCAGGATGGGG	TTGGCGCTCA	CGCGGTATCT	GAACAGCGAG	GGACGCTAGG	CGCGGCTTGA	540
TAATCTCGAG	AAAGCCATCG	CGTGGTGCCTA	GGAGCACATC	ACCTCGGGCG	ACCCCACTT	600
CCAGTTGGAG	GTTCGGGACAG	TCTACAACTC	GCTGTACAAAC	CGGAAAGGGG	AATACCGAGTC	660
ACTAGACTT	CGCTTTCCAT	ATTCGGATGC	GTTCGGAT	GTGGTGTTC	TTACCTCGGT	720
GTRCAACCCAC	ATGTTTCCCG	CGGACGCTGA	GCATCTATCTG	GACGAGATCT	CCCGCGTGTG	780
GAAGCCCGGC	GGACGATGTC	TCTTCACACTA	CTTCTGGCTC	ATGACGAGGT	CGTIAAGCCA	840
CTTCGGGAA	GGAAAGAGTG	CGCACAACCTT	CGGCGATGAG	GGACCGGGTT	ATCGGACAT	900
CCACGAAAG	GGGGCGGAG	AAGCAATCGG	CTTGGCGGAG	ACCTTCGTC	GGGATGTC	960
TGGCAAGTT	GGCTTCCCG	TCACCGAAC	ATTCGACTAC	GGCTCATGGA	GTGGCCGGGA	1020
ACCCAGCTTA	AGGTTTCAAGG	ACATCGTGT	CGGACGACAA	ACCGCGAGCT	AGGTCGGCAT	1080
CGGGGAAGCA	TGGCGACACC	GTGGCGCGGA	GCCTCCCTCG	GGGAGCGCG	ATTTAGCGGG	1140
CAGATTTGAC	GGGGCGGGGT	CGGCGCTCGG	AGTACCGGGC	CCCGAATGGC	GTCACCGGGCT	1200
GTTAACACCG	CTTGGCGGCC	TTGGGCGCGG	CTTGGCGGAT	CAGGTGGTAC	ATTCGACACA	1260
ACCCGCGTGT	ATCGGTCATC	ACCAACGCGT	ACACGACGGG	GTGGTGCACC	AGGGCGAACG	1320
CGACACCGGT	CTTCGGGCGG	TCACCGCGG	TOAAGCGCGG	CAAGCCAC	TGACCAAACC	1380
CGGGCATCAC	CGGGCGGATC	GGCATACCG	GATGGCGAC	ATGAAATT	AAGGGACCCA	1440
ATAGATTTGC	ATCCCGCAGA	ACTTGGCGTC	GGTGGCGGT	CAGGGCGTGT	ACCAAGCTCC	1500
GGGACAAGAA	CGTATGCGG	TGATCTCGC	CTCTGGCG			1539

(2) INFORMATION FOR SEQ ID NO:33:

(1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 551 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:33:

CTTCAGGGTG	GGCTGGATGTA	GGCTCACCGC	GGGCGAGCGC	GAGCTGACCG	GGCCCCAGGT	60
CGGGGGTGT	GGGGCGGCGCT	ACCGAGACGGT	GTATGGCTG	ACGGTGCCCC	CGCGGGGTGAT	120
GGCGGAGAAC	CGTCTGTGAC	GTATGGATTC	GTAGCGACG	ACCTCTTGG	GGCGAACAC	180
CCCGCGGATC	GGGGTCAACG	ASGGCGATAA	CGGGCGAGAT	TGGGGCGCG	AGCGGGCGCG	240
GATGTTTGGC	TACGGCGCGG	CGACCGCGAC	GGGACGCGCG	ACCTTGCTGC	CGTTGGAGGA	300
GGGGCGGGAG	ATGGACCGGG	GGGGTGGGGT	CTTCGACGAG	GGCGCGCGG	TGAGGAGGC	360
CTTCGGACAC	GGGGCGGGGA	ACAGGTGAT	GAACATGTC	CCCGAGCGCG	TGAAACAGTT	420
GGCGGGACCC	ACGGCGGGG	CCACCGCTC	TCCGAGCTG	GGTGGCGCTG	GGAAAGACGGT	480
CTTCGGCGAT	CGTGGCGCGA	TCAGCAACAT	GGTGGCGATG	GGCAACACCC	ACGTTGGAT	540
GACCGACTCG	GGTGTGCGA	TCACCGAACAC	CTTGAGCTG	AGTTGAAAGG	GCTTGGCTG	600
GGGGCGGGCG	GGGGCGGGCG	TGCAAAACAC	GGGGCGGGCG	GGGGCGGGCG	CGATGAGCTC	660
GCTGGGGCGC	GGGGCGGGGT	CTTCGGGGGT	GGGGCGGGGG	GTGGCGGCCA	ACTTGGGTG	720
GGGGCGGGCTG	GTACGGGTATG	GTCACTGGGA	GGGGCGGGAA	TATGCAAGATG	CTGGTGGCG	780
GAACGGTGGT	GGGGCGTAAAG	GTTCACCCCC	GGTGGCTGG	TGCGTGAAC	TTGCTCAACG	840
GAACACGTTA	C					851

(2) INFORMATION FOR SEQ ID NO:34:

(1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 254 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:34:

GATCGATCGG GCGGAAATTG GGACCAAGATT CGCCCTCGGC GATRACCCAA TCAATCGAAC	50
CTAGATTTAT TCGGTCCAGG GCGCCGAGTA ATGGCTCGCA GGAGAGGAAC CTTACCTGCTG	120
CGGGCGCTG TCGTAGGTCG TCGATAGGGC GGAGGGCGTC GACATTTCG ACCGAGACCC	180
CGATCGAAAC GTTGGAGGGC CACTCCAGCT TGTGAGGAG GCGACGCGAGT CGCAGGGTCT	240
GCTTGGTCAA GATC	284

(2) INFORMATION FOR SEQ ID NO:35:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1227 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) SEQUENCE DESCRIPTION: SEQ ID NO:35:

GATCTGACG GAAAGCTCGG CGGCCNAGGC GAACTCGCTG TTGGACCAGG AGGGACGGGA	63
CGATCTGGG CGTCCGGATCG CGCTTCAGGC GGGGGGGGTCG GCTGGATTTCG GCTATACCT	123
TTTCTTCGAC GACCCGGAGC TCGATGTCG CCACACCGCC TGGTCCGTT GTGTCAGGTT	183
GATGTTGGAC CGGATGAGGCG CGCGTATGATG GGGAGGCGG TGATGCTATT TGGTTCACAC	243
TATTGAGGAG CGAAGTTAC CATTCGACAT CGCAAGCGCA CGCGCTCTG CGCTGCGGG	303
CGATTCCTTA ACTGATTTAA CGCTAGTCG ACCCGCGATG CGCAGACACG TACCGACAC	363
CCAGGACCTG ACCCGCTGCA AAAGCAACT GAGGGATGCC TGGCACTTGA CGGGTGGCG	423
GGCCCGCCCGG GGCAGGTGTC ACCCTGCGG TGACGACAC GCGCTGCTG TATTGGGAC	483
AGTACACGAT TTGTCGATC GAGGTCACT CGACCTGGGA GAACTGCTTG CGGAACCGCT	543
CGCTGCTCG CTTGGGAAAG CGCTATCG CGCGCTTTCG GCGCACCGCG TGCTGGATAC	603
CGCACAGCGC ATTCGACACG ATGGCTGTCG CATCGCGTTT CTCCAGGGCG TTGGAGGTAC	663
CGTGAATGCG GGTCTGGGG CGTCCCTCG AGATGATGCC TGGCGCTGTT GCTCGCTTGG	723
TGGGACCCCC GTATATGATC CGCCCGCTCA TAGCGACACG CAGCGCGAGG GCTACCAAC	783
TGGCGATCG CAGCGCGCTG TCGCTCGCT TCGGGTAGGA CACCTTGCGC GGACACCGCG	843
GATATGCGGC GGCGCGCGAG CGCGCGCTTC CTGCGCTGTT CGGGCGAGAG CGCGCTTGG	903
CGGCGCGCGAG CTGCTGGGGG TAGTCGAGG CTGGCGCTTC CTGGGATAGG GCGCTCGCT	963
ACGCGCGCGAG TCGCTGGCTG CGGACACCGG GGTTCGGGGCA GTGGGGACCG CGCATTTGG	1023
TTTCTCTGGG GTGGTGGAGG CGGACACCGG CTGAGGGCGAC AACCGCGCGT CGGGTGGCG	1083
GGCACCGATCG GCGATCGAGT GAGCTTCTTA GGCGAGCTTG CGCACACGCG GCGCTCGCT	1143
CTCAACCGCA CGGGGGGGCG CGCGCGCGCG ATTAATGTTGA AGAGCTGGC AACCTTGGAG	1203
ACGAAAGACG GAGATTTTGT GACGATC	1227

(2) INFORMATION FOR SEQ ID NO:36:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 181 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) SEQUENCE DESCRIPTION: SEQ ID NO:36:

GGGGTGTCCG CGGATCGGGC GGGTGGTTGA AGCGCAACAGG CGGGGGCGGC GGGGGCGCG	63
GGACCGCGCG CGTACGGTGTG GCGGGGGCGA AGCGCTGTTT GTTGGGGGCC GGGGGGCTCG	123
GGCGCGCGCG CACCGATGTT GGTGTCGGCG GGTTCGGGCC ATTTCGTCAC CGCAACCGCG	183
G	181

(2) INFORMATION FOR SEQ ID NO:37:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 290 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:37:

GGCGGTCTCG	CGGATCCGGC	GGGTGGTTGA	ACGGCAACGG	GGGTGCGC	GGCGGGGGCG	60
GCGACGGCGT	CTTTCGGCGT	GGCGCGGCC	AGGGCGGCCT	GGGTGGCGAG	GGCGGCAATG	120
GGGGCGGGCT	CACCGCGCGC	ACCGCGCGTC	TGGCGGGCGC	GGGCCTGGCG	GGAGGCGAACG	180
CCCGCGGACG	CGGCTTGGGT	GGCACCGGCG	GTAAGGGTGC	CCAGGGCGGN	ATTCGGCGCG	240
GCACATCAGAG	CGGGACCGGC	CTGGGNGGTG	ACGGGGGTGA	GGGGCGTCAC		290

(2) INFORMATION FOR SEQ ID NO:38:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 34 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:38:

GATCCAGTGG	CATGGGGGGT	GTCAGTGGAA	GCAT		34
------------	------------	------------	------	--	----

(2) INFORMATION FOR SEQ ID NO:39:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 158 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:39:

GATCGCTCT	CTGTCACCCCG	TTGGCGCGGA	CGCCACCGGT	CCGACCGTTA	CCGAACACAGC	60
TGGCTTGCTG	GCACACACCG	CGGACACCG	CGACCGCGGA	GTCGAAACAT	GGCACCGTGTG	120
TATCCCGGAC	ATTCGGCGCC	GGCGACCCCG	CACCG			185

(2) INFORMATION FOR SEQ ID NO:40:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 83 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:40:

ATGGGGTTCA	CGGGGGCGCG	GGGACCGGGC	AGCCCGGNGG	GGCCGGGGGG	TGG	53
------------	------------	------------	------------	------------	-----	----

(2) INFORMATION FOR SEQ ID NO:41:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 132 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:41:

GATCCACCGC CGGTGTCAGAC GGTGCCCCGGC CGGGCACCCG GACCAAGGGC GCGAACGGCG	60
GCACCGGGCGG CAACGGGGCGG AACGCCACCGC TCGTGGGGG GGCCTGGGGG GCGGGCGGCA	120
AGGGCGGGCA CG	132

(2) INFORMATION FOR SEQ ID NO:42:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 132 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:42:

GATGGCCGGC CGGNACGGNC CGGGACGGCG GCGAGGGCGG AAACGGGGCG GCGGAGCCA	60
CGGGCGGAAAGA ATCCCTGGG TGCGGCAATG GCGCGAATGG CGGACAGGGC GGCAACGGCG	120
GCANCGGGCGG CA	132

(2) INFORMATION FOR SEQ ID NO:43:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 702 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:43:

CGGCACGAGG ATCGTGTACCC CGCRCATCGC GCAGCTGGCG ATTCCCGGGG TTTCCCCACC	60
CGAGGAAAGC CGCTTACCAAGA TGCGCTCGGC GAAGTAGGGC GATCGGTTGG CGATGCCGGC	120
ATGAAACGGGC CGCGAAATAT TACTCGAAGA ACCTTGTAGT TTAGCGACCA TAATGGCTAT	180
AGCACTAAGG AGGATGATCC GATAGGACCC AGTCGAGACG CGTGCACGGTG GATCACAGAG	240
AGATTTTTGAA CGGGCGAAC GAGGTGGGGG CGCGATGGC GGACCCACCG ACTGTATGTC	300
CGATCACACC GTCGGAACTC AGCGGGGNTA AAACCGCOSC CGAACAGNTG GTTTTGCTCG	360
CGGACAACAT CGGGGAAATAC CTGGCGCGGC GTGCCAAAGA GCGGCGAGGT CTGGCGACCT	420
CGCTCGCCAA CGGGCGCCAG ATGTTATGGCG AGTTGATGAA GGAGGCTGGC ACCGGCTCGG	480
ACGAGGACGG CGAGGAACT GTGGAGGGAG AATGGCGGGG GGCGCTCGGA GGGGACAGTT	540
CGGGCGAACT AACCGTACCG CGGAGGGTGG CGACGGCGGG TGAAACCCRAAC TTCAATGGTC	600
TCAAGGAAGC GGCAAGGGAG CTGGAAACCGG GCGAACCAAGG CGCATCGTC CGCGACTGG	660
GGGATGGGTG GAAACTTTC ACCTGTGAGCGC TCGAGGGCGA CG	722

(2) INFORMATION FOR SEQ ID NO:44:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 198 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single

(xi) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:44:

GAAGCCGCGAC	CCCTGTCGGG	CGACGTTGGG	GTCAAACCGG	CATCGCTGG	TGGCGGTGGA	60
GGCGCGCGGGG	TGCCUTGGC	GGCGTTGGGA	TCCCGATCG	GGGGCGCGGA	ATCGTCTCGG	120
CUCGCTGGCG	CTGGTGTACAT	TGCCGGCTTA	GGCCAAAGGA	GGGCCGGCGG	GGGGCGCGCG	180
CTGGCGGGG	GTGGCATGGG	AAIGCCGATG	GGTGGCCCGC	ATCAGGGACA	AGGGGGCGGC	240
AACTCCAAGG	GTTCCTACCA	GGAAACAGAG	GGCTCTTACA	CGAGGATOC	TGTTGCGG	298

(2) INFORMATION FOR SEQ ID NO:45:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1088 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:45:

CCGCACGAGG	ATGGAATGGC	GTGCGGGG	GCACAGCGTC	GCACGTGACCC	ATGAGGAGAG	60
CCATGACCTA	CTTCGGGGT	AACCGGGAT	ACCCCGAACG	GCAGCCCGCA	GGTCCCTACG	120
GAGGCCTCAC	ACCTCTCTTC	GGCACCGCGG	ATGAGGGTGC	GAGCGAAGCTA	CGGATGATCC	180
TSAAACATCGC	CTGGCGCTG	CTGGCTCTG	CTGGTACTCT	GGCCAGCTTC	GGCCCATATG	240
TCACCTCTCG	TCGGCACTG	GGGGGGGGTG	ATGGCGCGAGT	GTCCCGTGAC	ACTGGGGTGC	300
CGCTGGGGGT	GGCTCTCTG	GTGCGCTCG	TTGCCCGGGT	GGTTCTGTT	CTAAAGGCA	360
AGAGCCATGT	GAGCTGATG	GGCTGATGTT	GGGTAATCGG	CTATTTCTG	ATGUTCTCGG	420
CGACCTTTAA	CAAGCCCAAG	GCCTATTGGA	CGCGTGGGGC	ATTGTTGGTT	GTGTTGGCTT	480
TCATCTGTGT	CCAGGGGTT	GGGGGAGTC	TTGGCGCTCT	GGTGGAGACC	GGCGCTATCA	540
CGCGCGCGGC	CGGGCGCCOC	ATGTTGGAC	CGTATGGACA	GTACGGGGGG	TACGGGCACT	600
ACCGGGCACTA	CGGGGGTCAAG	GGGGTTGGGT	ATACGGCTA	GCAGGGTGT	CACGAGGCGG	660
CGGAGCTACA	GTGCGCGCGC	CGCGAGCAGT	CTGGCGAGCT	TCGGGATAT	GGTCCCACT	720
AGCGGCGCTA	TTGGCTCTG	CGGAGCGAT	CGGGCAGTGG	ATACAGCT	CAAGCCCCGG	780
CTTACCGCGGC	GGCGCGCTGC	ATGGCGCAC	GGGCTCATCC	ACGGCACCTA	840	
CGGGCTTTCG	GAAGCTTCAGC	CGGACACAC	CGTCACTGC	CGGGACGGGG	TGGCAGGCTG	900
GTTCGGCTCC	AGTCCTAACAT	TCGAAACGCC	GGGGGGGGA	CGAGTGTGCG	TCCCCCGGGG	960
GGGGCGCGGT	CTAACCGGGG	GTTCGGCGG	CGGGTGGCG	GTGTCGCGGA	AGAGTGAACA	1020
GGGTTGTCAGC	AAACGCGUGAC	GATCTCTGTC	CGGAAATTC			1088

(2) INFORMATION FOR SEQ ID NO:46:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 327 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:46:

GGGCGCGAGA	GACCGATGCC	GTCACTCTG	CGGAGGAGGC	AGGTAATTG	GACCGAGAT	60
CGGCGGACCT	GAAGACCGAC	ATGGACCCAGG	TTGGATGAGAC	GGCAGGTTG	TTGCAAGGGCC	120
AGTGGCGCGG	CGGGCGGGGG	ACGGCGCGG	AGGCGCGGGT	GGTGCCTTC	CGAGAGCG	180
CGGATGAGCA	GAAGCGGAA	CTCGACGAGA	TTGGCGAGGA	TATTCGTCAG	GGGGCGCTCC	240
AAATCTCGAG	GGCGGACGAG	GGGGCGGAG	AGGGCGCTTC	CTCGCAATG	GGTTCTGAC	300
CGGCGTAAAC	AAAAAGAAC	GGGCGCA				327

(2) INFORMATION FOR SEQ ID NO:47:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 170 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:47:

CGCTGCGAT GATGGCCTTG TCGAACGTGA CCGATTCCTGT ACCGGCGCTCG TTGAGATCAA
 CCAACAAACGT GTTGGCCTCG GCAAAATGTGC CGNAACCCGTC GATCTCGGTG ATCTTGTCT
 TCTTCATCG GAAGTTCACA CGGGCCACCC TGGCTCTCGN TACCTTTGCGG

60
120
170

(2) INFORMATION FOR SEQ ID NO:48:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 127 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:48:

GATCCGGCGG CACGGGGGGT GCGGGCGGCA GCACCGCTCG CGTGGCGGG AACGGGGGG
 CGGGGGCTGG CGGGGGAACT GGTGGGGTGC TCTTGGCAA CGGGGGTGCC CGGGGGGACG
 CGGGCGT

60
120
127

(2) INFORMATION FOR SEQ ID NO:49:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 81 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:49:

CGCGGGCAAG CGCGGCACCG CGCGCAACGG GAGCGGGCGG CGCGGGGGCA ACGGGGGCAA
 CGGGGGCTCG CGCTTCAACG G

60
81

(2) INFORMATION FOR SEQ ID NO:50:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 149 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:50:

GATCAGGGCT GCGGGGTCTC GCGCAGAAGG GCGGTACGG AGGAGCTGGG GGATTGTTTG
 GCAACGGCGG GCGGGGGGGT GCGGGGGGT CCAACCAAGC CGGTACGGG CGGGGGGGCG
 GAAACGGTGG TGGCGGTGGG CTGATCTGG

60
120
149

(2) INFORMATION FOR SEQ ID NO:51:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 355 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:51:

```

CGGCAAGAGA TCACACCTAC CGGTGATGG AGATGCTGG GACCTGGCCG GACGCTGTCG      60
ACCGGGAAT CGAGGGCGT CTGGCGCGAG CTGGCGCACG CATGCGCGG CTGGGACTGT      120
TCCAAGTACA TGCAATTGGA GGGCACCTGG TCGACGGAGC GGTGCGCAC TCCAGGTTGA      180
CTATGAAAGT CGGCTTCCG CTGGAGGTT CTGAGACCTT CAAGCGGGC CGATAACTGA      240
GTTGCGATCAT TAAGGACTT TTCCAGAGCA TCTGACGGG CTCGAAACCG GGTTCAGGCG      300
ACCGTGGCTC CGCGAGGCG CTGGCTCCAA AATCGTGGG ACAAATCGTC GCGCGG      355

```

(2) INFORMATION FOR SEQ ID NO:52:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 999 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:52:

```

ATGCGATCAG ATCACCATCA CAGTCATCA GTCGGACCTCA ACTTGACACG TCGCAAGGGA      60
CGATTTGGGG CACTGGCTAT CGCGCGCATG CGCGCGCCG GCGTGGTGTAC CGTTGGGGTG      120
CGCGCGACCG CGAACCGCGA TCGGGAGGCC CGCGCGCGCG TACCCGACAC CGCCGCGCTCG      180
CGCGCGCTGA CGCGCGACCG CGAACCGCGA CGCGCGACAC CGTGGCGCC CGCACACCG      240
CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG CGCGCGCGCG      300
GACGACCGG TTGGAGGATT CGAGCTTGCG CTGGCTGGTG GCGGGGGTGGG GTCTGACGCGC      360
GCCCGACTTGC ACTGGCTTC AGGACTCTTC AGCGAAACCA CGGGGGACCC GCCAITTCCC      420
GGACGACCGG CGCGCGCTGGC CAAGGACACC CGTATGCTGC TCGGCCTGGT AGACCAAAG      480
CTTGGACCCCA CGCGCGACCG CGCGGACTTC AGGGCGCGG CGCGGGTTGGG CTGGACATCG      540
GGTGGAGTTT ATGAGCTTC CGCGCGACCG CGCGGACCC AGGAAACCGT CTGCGCTGAC      600
GCCAACCGGG TTGGAGGAG CGCGCTGAT TACGAAGTCA AGTTGACGGA TCGGAGTAAAG      660
CGGACGCGCC AGATGTGGAC CGCGCTAATG CGCTCGCCCG CGCGGAAACCG ACCGGGACCGC      720
GGGCGCGCTT AGCGCTGTT TGTGGTATGG CGCGGGACCG CGAACAAACCC GGTGGACACAG      780
GGGCGCGCGCA AGGCGCTGGC CGAATGCTGC CGCGCTTGGG TCGCCCGCC CGCGGGCGCG      840
CGAACCGGCTC CTGCGAGACG CGCGTGGGG CGCGGGCGCG CGCGGGAAAGT CGCTGCTAAC      900
CGAACACAGC CGAACACCGA CGCGACCTTA CGCGCGTGA      960

```

(2) INFORMATION FOR SEQ ID NO:53:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 312 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:53:

Met His His His His His His Met His Gln Val Asp Pro Asn Leu Thr
 1 5 10 15
 Arg Arg Lys Gly Arg Leu Ala Ala Leu Ala Ile Ala Asn Met Ala Ser
 20 25 30
 Ala Ser Leu Val Thr Val Ala Val Pro Ala Thr Ala Asn Ala Asp Pro
 35 40 45
 Glu Pro Ala Pro Pro Val Pro Thr Thr Ala Ala Ser Pro Pro Ser Thr
 50 55 60
 Ala Ala Ala Pro Pro Ala Pro Ala Thr Pro Val Ala Pro Pro Pro Pro
 65 70 75 80
 Ala Ala Ala Asn Thr Pro Asn Ala Gln Pro Gly Asp Pro Asn Ala Ala
 85 90 95
 Pro Pro Pro Ala Asp Pro Asn Ala Pro Pro Pro Pro Val Ile Ala Pro
 100 105 110
 Asn Ala Pro Gln Pro Val Arg Ile Asp Asn Pro Val Gly Gly Phe Ser
 115 120 125
 Phe Ala Leu Pro Ala Gly Tyr Val Glu Ser Asp Ala Ala His Phe Asp
 130 135 140
 Tyr Gly Ser Ala Leu Leu Ser Lys Thr Thr Gly Asp Pro Pro Phe Pro
 145 150 155 160
 Gly Gln Pro Pro Pro Val Ala Asn Asp Thr Arg Ile Val Leu Gly Arg
 165 170 175
 Leu Asp Gln Lys Leu Tyr Ala Ser Ala Glu Ala Thr Asp Ser Lys Ala
 180 185 190
 Ala Ala Arg Leu Gly Ser Asp Met Gly Glu Phe Tyr Met Pro Tyr Pro
 195 200 205
 Gly Thr Arg Ile Asn Gln Glu Thr Val Ser Leu Asp Ala Asn Gly Val
 210 215 220
 Ser Gly Ser Ala Ser Tyr Tyr Glu Val Lys Phe Ser Asp Pro Ser Lys
 225 230 235 240
 Pro Asn Gly Gln Ile Tyr Thr Gly Val Ile Gly Ser Pro Ala Ala Asn
 245 250 255
 Ala Pro Asp Ala Gly Pro Pro Gln Arg Trp Phe Val Val Trp Leu Gly
 260 265 270
 Thr Ala Asn Asn Pro Val Asp Lys Gly Ala Ala Lys Ala Leu Ala Glu
 275 280 285
 Ser Ile Arg Pro Leu Val Ala Pro Pro Pro Ala Pro Ala Pro Ala Pro
 290 295 300
 Ala Glu Pro Ala Pro Ala Pro Ala Pro Ala Gly Glu Val Ala Pro Thr
 305 310 315 320
 Pro Thr Thr Pro Thr Pro Gln Arg Thr Leu Pro Ala
 325 330

(2) INFORMATION FOR SEQ ID NO:54:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

(ii) SEQUENCE DESCRIPTION: SEQ ID NO:54:

Asp Pro Val Asp Ala Val Ile Asn Thr Thr Lys Asn Tyr Gly Gln Val

1	5	10	15
Val	Ala	Ala	Leu
		20	

(2) INFORMATION FOR SEQ ID NO:55:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 15 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:55:

Ala	Val	Glu	Ser	Gly	Met	Leu	Ala	Leu	Gly	Thr	Pro	Ala	Pro	Ser
1					5					10				15

(2) INFORMATION FOR SEQ ID NO:56:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 19 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:56:

Ala	Ala	Met	Lys	Pro	Arg	Thr	Gly	Asp	Gly	Pro	Leu	Glu	Ala	Ala	Lys
1					5					10				15	
Glu	Gly	Arg													

(2) INFORMATION FOR SEQ ID NO:57:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 15 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:57:

Tyr	Tyr	Trp	Cys	Pro	Gly	Gln	Pro	Phe	Asp	Pro	Ala	Trp	Gly	Pro
1					5					10				15

(2) INFORMATION FOR SEQ ID NO:58:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 14 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:58:

(2) INFORMATION FOR SEQ ID NO: 20

(1) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 11 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS:
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 60

Ala Glu Glu Ser Ile Ser Thr Xaa Xaa Glu Xaa Ile Val Pro
1 5 10

(2) INFORMATION FOR SDO IN AREA 50

(1) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 17 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(iii) SEQUENCE DESCRIPTION: SEQ ID NO: 2

Asp Pro Glu Pro Ala Pro Pro Val Pro Thr Ala Ala Ala Ala Pro Pro
 1 S 10 15
 Ala

(2) INFORMATION FOR SPC IN AND OUT

(4) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 15 amino acids
- (B) TYPE: amino acid
- (C) STRANGENESS:
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: [seq](#) [src](#)

Ala Pro Lys Thr Tyr Xaa Glu Glu Leu Lys Gly Thr Asp Thr Gly
1 5 10 15

202 INFORMATION FOR CBO IN 100-100

(1) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 30 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTIONS - 2000

Asp Pro Ala Ser Ala Pro Asp Val Pro Thr Ala Ala Gln Gln Thr Ser
 1 5 10 15

Leu	Leu	Asn	Asn
Leu	Ala	Asp	Pro
26	26	26	26
Asp	Val	Ser	Phe
26	26	26	26
Val	Asp	Ala	Asp
26	26	26	26

(2) INFORMATION FOR SEQ ID NO:63:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 187 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:63:

Thr	Gly	Ser	Leu	Asn	Gln	Thr	His	Asn	Arg	Arg	Ala	Asn	Glu	Arg	Lys
1	5	9	13	17	21	25	29	33	37	41	45	49	53	57	61
Asn	Thr	Thr	Met	Lys	Met	Val	Lys	Ser	Ile	Ala	Ala	Gly	Leu	Thr	Ala
20	24	28	32	36	40	44	48	52	56	60	64	68	72	76	80
Ala	Ala	Ala	Ile	Gly	Ala	Ala	Ala	Ala	Gly	Val	Thr	Ser	Ile	Met	Ala
35	39	43	47	51	55	59	63	67	71	75	79	83	87	91	95
Gly	Gly	Pro	Val	Val	Tyr	Gln	Met	Gln	Pro	Val	Val	Phe	Gly	Ala	Pro
90	94	98	102	106	110	114	118	122	126	130	134	138	142	146	150
Leu	Pro	Leu	Asp	Pro	Ala	Ser	Ala	Pro	Asp	Val	Pro	Thr	Ala	Ala	Gln
65	70	75	80	85	90	95	100	105	110	115	120	125	130	135	140
Leu	Thr	Ser	Leu	Asn	Ser	Leu	Ala	Asp	Pro	Asn	Val	Ser	Phe	Ala	Asn
85	90	95	100	105	110	115	120	125	130	135	140	145	150	155	160
Asn	Lys	Gly	Ser	Leu	Val	Glu	Gly	Ile	Gly	Gly	Thr	Glu	Ala	Arg	Lys
100	105	110	115	120	125	130	135	140	145	150	155	160	165	170	175
Ile	Ala	Asp	His	Lys	Leu	Lys	Ala	Ala	Glu	His	Gly	Asp	Leu	Pro	Leu
115	120	125	130	135	140	145	150	155	160	165	170	175	180	185	190
Leu	Ser	Phe	Ser	Val	Thr	Asn	Ile	Gln	Pro	Ala	Ala	Gly	Ser	Ala	Asn
130	135	140	145	150	155	160	165	170	175	180	185	190	195	200	205
Thr	Ala	Asp	Val	Ser	Val	Ser	Gly	Pro	Lys	Leu	Ser	Ser	Pro	Val	Thr
145	150	155	160	165	170	175	180	185	190	195	200	205	210	215	220
Gln	Asn	Val	Thr	Phe	Val	Asn	Gln	Gly	Gly	Trp	Met	Leu	Ser	Arg	Ala
165	170	175	180	185	190	195	200	205	210	215	220	225	230	235	240
Ser	Ala	Met	Glu	Leu	Leu	Gln	Ala	Ala	Gly	Xaa					

(2) INFORMATION FOR SEQ ID NO:64:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 148 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:64:

Asp	Glu	Val	Thr	Val	Glu	Thr	Ser	Val	Phe	Arg	Ala	Asp	Phe	Leu	
1	5	9	13	17	21	25	29	33	37	41	45	49	53	57	61
Ser	Glu	Leu	Asp	Ala	Pro	Ala	Gln	Ala	Gly	Thr	Glu	Ser	Ala	Val	Ser
20	24	28	32	36	40	44	48	52	56	60	64	68	72	76	80
Gly	Val	Glu	Gly	Leu	Pro	Pro	Gly	Ser	Ala	Leu	Leu	Val	Val	Lys	Arg
35	40	45	50	55	59	63	68	72	76	80	85	90	95	100	105
Gly	Pro	Asn	Ala	Gly	Ser	Arg	Phe	Leu	Leu	Asp	Gln	Ala	Ile	Thr	Ser

50	55	60
Ala Gly Arg His Pro Asp Ser Asp Ile Phe Leu Asp Asp Val Thr Val		
65	70	75
Ser Arg Arg His Ala Glu Phe Arg Leu Glu Asn Asn Glu Phe Asn Val		80
85	90	95
Val Asp Val Gly Ser Leu Asn Gly Thr Tyr Val Asn Arg Glu Pro Val		
100	105	110
Asp Ser Ala Val Leu Ala Asn Gly Asp Glu Val Gln Ile Gly Lys Leu		
115	120	125
Arg Leu Val Phe Leu Thr Gly Pro Lys Gln Gly Glu Asp Asp Gly Ser		
130	135	140
Thr Gly Gly Pro		
145		

(2) INFORMATION FOR SEQ ID NO:65:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 210 amino acids
- (B) TYPE: amino acid
- (C) STRANGENESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:65:

Thr Ser Asn Arg Pro Ala Arg Arg Gly Arg Arg Ala Pro Arg Asp Thr			
1	3	10	15
Gly Pro Asp Arg Ser Ala Ser Leu Ser Leu Val Arg His Arg Arg Gln			
20	25	30	
Gln Arg Asp Ala Leu Cys Leu Ser Ser Thr Gln Ile Ser Arg Gln Ser			
35	40	45	
Asn Leu Pro Pro Ala Ala Gly Gly Ala Ala Asn Tyr Ser Arg Arg Asn			
50	55	60	
Phe Asp Val Arg Ile Lys Ile Phe Met Leu Val Thr Ala Val Val Leu			
65	70	75	80
Leu Cys Cys Ser Gly Val Ala Thr Ala Ala Pro Lys Thr Tyr Cys Glu			
85	90	95	
Glu Leu Lys Gly Thr Asp Thr Gly Gln Ala Cys Gln Ile Gln Met Ser			
100	105	110	
Asp Pro Ala Tyr Asn Ile Asn Ile Ser Leu Pro Ser Tyr Tyr Pro Asp			
115	120	125	
Gln Lys Ser Leu Glu Asn Tyr Ile Ala Gln Thr Arg Asp Lys Phe Leu			
130	135	140	
Ser Ala Ala Thr Ser Ser Thr Pro Arg Glu Ala Pro Tyr Glu Leu Asn			
145	150	155	160
Ile Thr Ser Ala Thr Tyr Gln Ser Ala Ile Pro Pro Arg Gly Thr Gln			
165	170	175	
Ala Val Val Leu Xaa Val Tyr His Asn Ala Gly Gly Thr His Pro Thr			
180	185	190	
Thr Thr Tyr Lys Ala Phe Asp Trp Asp Gln Ala Tyr Arg Lys Pro Ile			
195	200	205	
Thr Tyr Asp Thr Leu Trp Gln Ala Asp Thr Asp Pro Leu Pro Val Val			
210	215	220	
Pha Pro Ile Val Ala Arg			
225	230		

(2) INFORMATION FOR SRO TO INDEX

1.2 SEQUENCE CHARACTERISTICS

(A) LENGTH: 132 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(iii) SEQUENCE DESCRIPTION: SEQ ID NO: 66.

Thr Ala Ala Ser Asp Asn Phe Gln Leu Ser Gls Gly Gly Gln Gly Phe
 1 5 10 15
 Ala Ile Pro Ile Gly Gln Ala Met Ala Ile Ala Gly Gln Ile Arg Ser
 20 25 30
 Gly Gly Ser Pro Thr Val His Ile Gly Pro Thr Ala Phe Leu Gly
 35 40 45
 Leu Gly Val Val Asp Asn Asn Gly Asn Gly Ala Arg Val Gln Arg Val
 50 55 60
 Val Gly Ser Ala Pro Ala Ala Ser Leu Gly Ile Ser Thr Gly Asp Val
 65 70 75 80
 Ile Thr Ala Val Asp Gly Ala Pro Ile Asn Ser Ala Thr Ala Met Ala
 85 90 95
 Asp Ala Leu Asn Gly His His Pro Gly Asp Val Ile Ser Val Asn Trp
 100 105 110
 Gln Thr Lys Ser Gly Gly Thr Arg Thr Gly Asn Val Thr Leu Ala Gln
 115 120 125
 Gly Pro Pro Ala
 130

(2) INFORMATION FOR SDO ID NO. 63

SEQUENCE CLASSIFICATION

SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 108 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 13

Val Pro Leu Arg Ser Pro Ser Met Ser Pro Ser Lys Cys Leu Ala Ala			
1	5	10	15
Ala Gln Arg Asn Pro Val Ile Arg Arg Arg Arg Leu Ser Asn Pro Pro			
20	25	30	
Pro Arg Lys Tyr Arg Ser Met Pro Ser Pro Ala Thr Ala Ser Ala Gly			
35	40	45	
Met Ala Arg Val Arg Arg Arg Ala Ile Trp Arg Gly Pro Ala Thr Xaa			
50	55	60	
Ser Ala Gly Met Ala Arg Val Arg Arg Trp Xaa Val Met Pro Xaa Val			
65	70	75	80
Ile Gln Ser Thr Xaa Ile Arg Xaa Xaa Gly Pro Phe Asp Asn Arg Gly			
85	90	95	
Ser Glu Arg Lys			
100			

12. INFORMATION FOR SEQ ID NO: 63

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 163 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:68:

```

Met Thr Asp Asp Ile Leu Leu Ile Asp Thr Asp Glu Arg Val Arg Thr
1           5           10          15
Leu Thr Leu Asn Arg Pro Gln Ser Arg Asn Ala Leu Ser Ala Ala Leu
20          25          30
Arg Asp Arg Phe Phe Ala Kaa Kaa Asp Ala Glu Kaa Asp Asp Asp
35          40          45
Ile Asp Val Val Ile Leu Thr Gly Ala Asp Pro Val Phe Cys Ala Gly
50          55          60
Leu Asp Leu Lys Val Ala Gly Arg Ala Asp Arg Ala Ala Gly His Leu
65          70          75          80
Thr Ala Val Gly Gly His Asp Gln Ala Gly Asp Arg Arg Asp Gln Arg
85          90          95
Arg Arg Gly His Arg Arg Ala Arg Thr Gly Ala Val Leu Arg His Pro
100         105         110
Asp Arg Leu Arg Ala Arg Pro Leu Arg Arg His Pro Arg Pro Gly Gly
115         120         125
Ala Ala Ala His Leu Gly Thr Gln Cys Val Leu Ala Ala Lys Gly Arg
130         135         140
His Arg Kaa Gly Pro Val Asp Glu Pro Asp Arg Arg Leu Pro Val Arg
145         150         155         160
Asp Arg Arg

```

(12) INFORMATION FOR SEQ ID NO:69:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 344 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:69:

```

Met Lys Phe Val Asn His Ile Glu Pro Val Ala Pro Arg Arg Ala Gly
1           5           10          15
Gly Ala Val Ala Glu Val Tyr Ala Glu Ala Arg Arg Glu Phe Gly Arg
20          25          30
Leu Pro Glu Pro Leu Ala Met Leu Ser Pro Asp Glu Gly Leu Leu Thr
35          40          45
Ala Gly Trp Ala Thr Leu Arg Glu Thr Leu Leu Val Gly Gln Val Pro
50          55          60
Arg Gly Arg Lys Glu Ala Val Ala Ala Val Ala Ala Ser Leu Arg
65          70          75          80
Cys Pro Trp Cys Val Asp Ala His Thr Thr Met Leu Tyr Ala Ala Gly
85          90          95
Gln Thr Asp Thr Ala Ala Ala Ile Leu Ala Gly Thr Ala Pro Ala Ala
100         105         110

```

Gly Asp Pro Asn Ala Pro Tyr Val Ala Trp Ala Ala Gly Thr Gly Thr
 115 120 125
 Pro Ala Gly Pro Pro Ala Pro Pro Phe Gly Pro Asp Val Ala Ala Glu Tyr
 130 135 140
 Leu Gly Thr Ala Val Gln Phe His Phe Ile Ala Arg Leu Val Leu Val
 145 150 155 160
 Leu Leu Asp Glu Thr Phe Leu Pro Gly Gly Pro Arg Ala Gln Gln Leu
 165 170 175
 Met Arg Arg Ala Gly Gly Leu Val Phe Ala Arg Lys Val Arg Ala Glu
 180 185 190
 His Arg Pro Gly Arg Ser Thr Arg Arg Leu Glu Pro Arg Thr Leu Pro
 195 200 205
 Asp Asp Leu Ala Trp Ala Thr Pro Ser Glu Pro Ile Ala Thr Ala Phe
 210 215 220
 Ala Ala Leu Ser His His Leu Asp Thr Ala Pro His Leu Pro Pro
 225 230 235 240
 Thr Arg Gln Val Val Arg Arg Val Val Gly Ser Trp His Gly Glu Pro
 245 250 255
 Met Pro Met Ser Ser Arg Trp Thr Asn Glu His Thr Ala Glu Leu Pro
 260 265 270
 Ala Asp Leu His Ala Pro Thr Arg Leu Ala Leu Leu Thr Gly Leu Ala
 275 280 285
 Pro His Gln Val Thr Asp Asp Asp Val Ala Ala Ala Arg Ser Leu Leu
 290 295 300
 Asp Thr Asp Ala Ala Leu Val Gly Ala Leu Ala Trp Ala Ala Phe Thr
 305 310 315 320
 Ala Ala Arg Arg Ile Gly Thr Trp Ile Gly Ala Ala Ala Glu Gly Gln
 325 330 335
 Val Ser Arg Gln Asp Pro Thr Gly
 340

(2) INFORMATION FOR SEQ ID NO:70:

(1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 485 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:70:

Asp Asp Pro Asp Met Pro Gly Thr Val Ala Lys Ala Val Ala Asp Ala
 1 5 10 15
 Leu Gly Arg Gly Ile Ala Pro Val Glu Asp Ile Gln Asp Cys Val Glu
 20 25 30
 Ala Arg Leu Gly Glu Ala Gly Leu Asp Asp Val Ala Arg Val Tyr Ile
 35 40 45
 Ile Tyr Arg Gln Arg Arg Ala Glu Leu Arg Thr Ala Lys Ala Leu Leu
 50 55 60
 Gly Val Arg Asp Glu Leu Lys Leu Ser Leu Ala Ala Val Thr Val Leu
 65 70 75 80
 Arg Glu Arg Tyr Leu Leu His Asp Glu Gln Gly Arg Pro Ala Glu Ser
 85 90 95
 Thr Gly Glu Leu Met Asp Arg Ser Ala Arg Cys Val Ala Ala Glu
 100 105 110

Asp Gln Tyr Glu Pro Gly Ser Ser Arg Arg Trp Ala Glu Arg Phe Ala
 115 120 125
 Thr Leu Leu Arg Asn Leu Glu Phe Leu Pro Asn Ser Pro Thr Leu Met
 130 135 140
 Asn Ser Gly Thr Asp Leu Gly Leu Leu Ala Gly Cys Phe Val Leu Pro
 145 150 155 160
 Ile Glu Asp Ser Leu Gln Ser Ile Phe Ala Thr Leu Gly Gln Ala Ala
 165 170 175
 Glu Leu Gln Arg Ala Gly Gly Thr Gly Tyr Ala Phe Ser His Leu
 180 185 190
 Arg Pro Ala Gly Asp Arg Val Ala Ser Thr Gly Gly Thr Ala Ser Gly
 195 200 205
 Pro Val Ser Phe Leu Arg Leu Tyr Asp Ser Ala Ala Gly Val Val Ser
 210 215 220
 Met Gly Gly Arg Arg Gly Ala Cys Met Ala Val Leu Asp Val Ser
 225 230 235 240
 His Pro Asp Ile Cys Asp Phe Val Thr Ala Lys Ala Glu Ser Pro Ser
 245 250 255
 Glu Leu Pro His Phe Asn Leu Ser Val Gly Val Thr Asp Ala Phe Leu
 260 265 270
 Arg Ala Val Glu Arg Asn Gly Leu His Arg Leu Val Asn Pro Arg Thr
 275 280 285
 Gly Lys Ile Val Ala Arg Met Pro Ala Ala Glu Leu Phe Asp Ala Ile
 290 295 300
 Cys Lys Ala Ala His Ala Gly Gly Asp Pro Gly Leu Val Phe Leu Asp
 305 310 315 320
 Thr Ile Asn Arg Ala Asn Pro Val Pro Gly Arg Gly Arg Ile Glu Ala
 325 330 335
 Thr Asn Pro Cys Gly Glu Val Pro Leu Leu Pro Tyr Glu Ser Cys Asn
 340 345 350
 Leu Gly Ser Ile Asn Leu Ala Arg Met Leu Ala Asp Gly Arg Val Asp
 355 360 365
 Trp Asp Arg Leu Glu Glu Val Ala Gly Val Ala Val Arg Phe Leu Asp
 370 375 380
 Asp Val Ile Asp Val Ser Arg Tyr Pro Phe Pro Glu Leu Gly Glu Ala
 385 390 395 400
 Ala Arg Ala Thr Arg Lys Ile Gly Leu Gly Val Met Gly Leu Ala Glu
 405 410 415
 Leu Leu Ala Ala Leu Gly Ile Pro Tyr Asp Ser Glu Glu Ala Val Arg
 420 425 430
 Leu Ala Thr Arg Leu Met Arg Arg Ile Glu Glu Ala Ala His Thr Ala
 435 440 445
 Ser Arg Arg Leu Ala Glu Glu Arg Gly Ala Phe Pro Ala Phe Thr Asp
 450 455 460
 Ser Arg Phe Ala Arg Ser Gly Pro Arg Arg Asn Ala Glu Val Thr Ser
 465 470 475 480
 Val Ala Pro Thr Gly
 485

(2) INFORMATION FOR SEQ ID NO:71:

(1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 267 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 12

Gly Val Ile Val Leu Asp Leu Glu Pro Arg Gly Pro Leu Pro Thr Glu
 1 5 10 15
 Ile Tyr Trp Arg Arg Arg Gly Leu Ala Leu Gly Ile Ala Val Val Val
 20 25 30
 Val Gly Ile Ala Val Ala Ile Val Ile Ala Phe Val Asp Ser Ser Ala
 35 40 45
 Gly Ala Lys Pro Val Ser Ala Asp Lys Pro Ala Ser Ala Gin Ser His
 50 55 60
 Pro Gly Ser Pro Ala Pro Gln Ala Pro Gln Pro Ala Gly Glu Thr Glu
 65 70 75 80
 Gly Asn Ala Ala Ala Ala Pro Pro Gln Gly Glu Asn Pro Glu Thr Pro
 85 90 95
 Thr Pro Thr Ala Ala Val Glu Pro Pro Val Leu Lys Glu Gly Asp
 100 105 110
 Asp Cys Pro Asp Ser Thr Leu Ala Val Lys Gly Leu Thr Asn Ala Pro
 115 120 125
 Glu Tyr Tyr Val Gly Asp Gln Pro Lys Phe Thr Met Val Val Thr Asn
 130 135 140
 Ile Gly Leu Val Ser Cys Lys Arg Asp Val Glu Ala Ala Val Leu Ala
 145 150 155 160
 Ala Tyr Val Tyr Ser Leu Asp Asn Lys Arg Leu Trp Ser Asn Leu Asp
 165 170 175
 Cys Ala Pro Ser Asn Glu Thr Leu Val Lys Thr Phe Ser Pro Gly Glu
 180 185 190
 Glu Val Thr Thr Ala Val Thr Trp Thr Gly Met Gly Ser Ala Pro Arg
 195 200 205
 Cys Pro Leu Pro Arg Pro Ala Ile Gly Pro Gly Thr Tyr Asn Leu Val
 210 215 220
 Val Gln Leu Gly Asn Leu Arg Ser Leu Pro Val Pro Phe Ile Leu Asn
 225 230 235 240
 Glu Pro Pro Pro Pro Pro Gly Pro Val Pro Ala Pro Gly Pro Ala Glu
 245 250 255
 Ala Pro Pro Pro Glu Ser Pro Ala Gln Gly Gly
 260 265

(2) INFORMATION FOR SPC IN NO: 32.

1428 S. S. TURKMAN, C. W. J. GRANGER AND R. H. HOGG

SEQUENCE CHARACTERISTICS:

- LENGTH: 97 amino acids
- TYPE: amino acid
- STRANDEDNESS: single
- TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 22

Leu Ile Ser Thr Gly Lys Ala Ser His Ala Ser Leu Gly Val Gln Val
 1 5 10 15
 Thr Asn Asp Lys Asp Thr Pro Gly Ala Lys Ile Val Glu Val Val Ala
 20 25 30
 Gly Gly Ala Ala Ala Asn Ala Gly Val Pro Lys Gly Val Val Val Thr
 35 40 45

Lys Val Asp Asp Arg Pro Ile Asn Ser Ala Asp Ala Leu Val Ala Ala
 50 55 60
 Val Arg Ser Lys Ala Pro Gly Ala Thr Val Ala Leu Thr Phe Gln Asp
 65 70 75 80
 Pro Ser Gly Gly Ser Arg Thr Val Gln Val Thr Leu Gly Lys Ala Gln
 85 90 95
 Gln

(2) INFORMATION FOR SEQ ID NO:73:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 364 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:73:

Gly Ala Ala Val Ser Leu Leu Ala Ala Gly Thr Leu Val Leu Thr Ala
 1 5 10 15
 Cys Gly Gly Gly Thr Asn Ser Ser Ser Ser Gly Ala Gly Gly Thr Ser
 20 25 30
 Gly Ser Val His Cys Gly Gly Lys Lys Glu Leu His Ser Ser Gly Ser
 35 40 45
 Thr Ala Gln Glu Asn Ala Met Glu Gln Phe Val Tyr Ala Tyr Val Arg
 50 55 60
 Ser Cys Pro Gly Tyr Thr Leu Asp Tyr Asn Ala Asn Gly Ser Gly Ala
 65 70 75 80
 Gly Val Thr Gln Phe Leu Asn Asn Glu Thr Asp Phe Ala Gly Ser Asp
 85 90 95
 Val Pro Leu Asn Pro Ser Thr Gly Gln Pro Asp Arg Ser Ala Glu Arg
 100 105 110
 Cys Gly Ser Pro Ala Trp Asp Leu Pro Thr Val Phe Gly Pro Ile Ala
 115 120 125
 Ile Thr Tyr Asn Ile Lys Gly Val Ser Thr Leu Asn Leu Asp Gly Pro
 130 135 140
 Thr Thr Ala Lys Ile Phe Asn Gly Thr Ile Thr Val Trp Asn Asp Pro
 145 150 155 160
 Gln Ile Gln Ala Leu Asn Ser Gly Thr Asp Leu Pro Pro Thr Pro Ile
 165 170 175
 Ser Val Ile Phe Arg Ser Asp Lys Ser Gly Thr Ser Asp Asn Phe Gln
 180 185 190
 Lys Tyr Leu Asp Gly Val Ser Asn Gly Ala Trp Gly Lys Gly Ala Ser
 195 200 205
 Glu Thr Phe Ser Gly Gly Val Gly Val Gly Ala Ser Gly Asn Asn Gly
 210 215 220
 Thr Ser Ala Leu Leu Gln Thr Thr Asp Gly Ser Ile Thr Tyr Asn Glu
 225 230 235 240
 Trp Ser Phe Ala Val Gly Lys Gln Leu Asn Met Ala Gln Ile Ile Thr
 245 250 255
 Ser Ala Gly Pro Asp Pro Val Ala Ile Thr Thr Glu Ser Val Gly Lys
 260 265 270
 Thr Ile Ala Gly Ala Lys Ile Met Gly Gln Gly Asn Asp Leu Val Leu
 275 280 285
 Asp Thr Ser Ser Phe Tyr Arg Pro Thr Gln Pro Gly Ser Tyr Pro Ile

(2) INFORMATION FOR SEO IN MD-34.

3.2 SEQUENCE CHARACTERISTICS

- (A) LENGTH: 309 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 34.

Gin Ala Ala Ala Gly Arg Ala Val Arg Arg Thr Gly His Ala Glu Asp
 1 S 10 15
 Gin Thr His Gln Asp Arg Leu His His Gly Cys Arg Arg Ala Ala Val
 20 25 30
 Val Val Arg Gly Asp Arg Ala Ser Val Ser Ala Thr Ser Ala Arg Pro
 35 40 45
 Pro Arg Arg His Pro Ala Gln Gly His Arg Arg Arg Val Ala Pro Ser
 50 55 60
 Gly Gly Arg Arg Arg Pro His Pro His His Val Gln Pro Asp Asp Arg
 65 70 75 80
 Arg Asp Arg Pro Ala Leu Leu Asp Arg Thr Gln Pro Ala Gln His Pro
 95 100 105 110
 Asp Pro His Arg Arg Gly Pro Ala Asp Pro Gly Arg Val Arg Gly Arg
 115 120 125
 Gly Arg Leu Arg Arg Val Asp Asp Gly Arg Leu Gln Pro Asp Arg Asp
 130 135 140 145
 Ala Asp His Gly Ala Pro Val Arg Gly Arg Gly Pro His Arg Gly Val
 150 155 160
 Gin His Arg Gly Gly Pro Val Phe Val Arg Arg Val Pro Gly Val Arg
 165 170 175 180
 Cys Ala His Arg Arg Gly His Arg Arg Val Ala Ala Pro Gly Gln Gly
 195 200 205
 Asp Val Leu Arg Ala Gly Leu Arg Val Glu Arg Leu Arg Pro Val Ala
 215 220 225
 Ala Val Glu Asn Leu His Arg Gly Ser Gin Arg Ala Asp Gly Arg Val
 235 240
 Phe Arg Pro Ile Arg Arg Gly Ala Arg Leu Pro Ala Arg Arg Ser Arg
 255 260 265
 Ala Gly Pro Gln Gly Arg Leu His Leu Asp Gly Ala Gly Pro Ser Pro
 275 280 285
 Leu Pro Ala Arg Ala Gly Gln Gln Gln Pro Ser Ser Ala Gly Gly Arg
 295 300 305
 Arg Ala Gly Gly Ala Glu Arg Ala Asp Pro Gly Gln Arg Gly Arg His
 320 325 330
 His Gln Gly Gly His Asp Pro Gly Arg Gln Gly Ala Gln Arg Ser Thr

275	280	285
Ala Gly Val Ala His Ala Ala Gly Pro Arg Arg	Ala Ala Val Arg	
290	295	300
Asn Arg Pro Arg Arg		
305		

(2) INFORMATION FOR SEQ ID NO:75:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 580 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:75:

Ser Ala Val Trp Cys Leu Asn Gly Phe Thr Gly Arg His Arg His Gly
 1 5 10 15
 Arg Cys Arg Val Arg Ala Ser Gly Trp Arg Ser Ser Asn Arg Trp Cys
 20 25 30
 Ser Thr Thr Ala Asp Cys Cys Ala Ser Lys Thr Pro Thr Gin Ala Ala
 35 40 45
 Ser Pro Leu Glu Arg Arg Phe Thr Cys Cys Ser Pro Ala Val Gly Cys
 50 55 60
 Arg Phe Arg Ser Phe Pro Val Arg Arg Leu Ala Leu Gly Ala Arg Thr
 65 70 75 80
 Ser Arg Thr Leu Gly Val Arg Arg Thr Leu Ser Gin Trp Asn Leu Ser
 85 90 95
 Pro Arg Ala Gln Pro Ser Cys Ala Val Thr Val Glu Ser His Thr His
 100 105 110
 Ala Ser Pro Arg Met Ala Lys Leu Ala Arg Val Val Gly Leu Val Gln
 115 120 125
 Glu Glu Gln Pro Ser Asp Met Thr Asn His Pro Arg Tyr Ser Pro Pro
 130 135 140
 Pro Gln Gln Pro Gly Thr Pro Gly Tyr Ala Gln Gly Gln Gln Gln Thr
 145 150 155 160
 Tyr Ser Gln Gln Phe Asp Trp Arg Tyr Pro Pro Ser Pro Pro Pro Gln
 165 170 175
 Pro Thr Gln Tyr Arg Gln Pro Tyr Glu Ala Leu Gly Gly Thr Arg Pro
 180 185 190
 Gly Leu Ile Pro Gly Val Ile Pro Thr Met Thr Pro Pro Pro Gly Met
 195 200 205
 Val Arg Gln Arg Pro Arg Ala Gly Met Leu Ala Ile Gly Ala Val Thr
 210 215 220
 Ile Ala Val Val Ser Ala Gly Ile Gly Gly Ala Ala Ser Leu Val
 225 230 235 240
 Gly Phe Asn Arg Ala Pro Ala Gly Pro Ser Gly Gly Pro Val Ala Ala
 245 250 255
 Ser Ala Ala Pro Ser Ile Pro Ala Ala Asn Met Pro Pro Gly Ser Val
 260 265 270
 Glu Gln Val Ala Ala Lys Val Val Pro Ser Val Val Met Leu Glu Thr
 275 280 285
 Asp Leu Gly Arg Gln Ser Glu Glu Gly Ser Gly Ile Ile Leu Ser Ala
 290 295 300
 Gln Gly Leu Ile Leu Thr Asn Asn His Val Ile Ala Ala Ala Lys

305	310	315	320
Pro Pro Leu Gly Ser Pro Pro Pro Pro Lys Thr Val Thr Ser Ser Asp			
325	330	335	
Gly Arg Thr Ala Pro Phe Thr Val Val Gly Ala Asp Pro Thr Ser Asp			
340	345	350	
Ile Ala Val Val Arg Val Gln Gly Val Ser Gly Leu Thr Pro Ile Ser			
355	360	365	
Leu Gly Ser Ser Ser Asp Leu Arg Val Gly Gln Pro Val Leu Ala Ile			
370	375	380	
Gly Ser Pro Leu Gly Leu Glu Gly Thr Val Thr Gly Ile Val Ser			
385	390	395	400
Ala Leu Asn Arg Pro Val Ser Thr Thr Gly Glu Ala Gly Asn Gln Asn			
405	410	415	
Thr Val Leu Asp Ala Ile Gln Thr Asp Ala Ala Ile Asn Pro Gly Asn			
420	425	430	
Ser Gly Gly Ala Leu Val Asn Asn Ala Gln Leu Val Gly Val Asn			
435	440	445	
Ser Ala Ile Ala Thr Leu Gly Ala Asp Ser Ala Asp Ala Gln Ser Gly			
450	455	460	
Ser Ile Gly Leu Gly Phe Ala Ile Pro Val Asp Gln Ala Lys Arg Ile			
465	470	475	480
Ala Asp Glu Leu Ile Ser Thr Gly Lys Ala Ser His Ala Ser Leu Gly			
485	490	495	
Val Gln Val Thr Asn Asp Lys Asp Thr Pro Gly Ala Lys Ile Val Glu			
500	505	510	
Val Val Ala Gly Gly Ala Ala Ala Asn Ala Gly Val Pro Lys Gly Val			
515	520	525	
Val Val Thr Lys Val Asp Asp Arg Pro Ile Asn Ser Ala Asp Ala Leu			
530	535	540	
Val Ala Ala Val Arg Ser Lys Ala Pro Gly Ala Thr Val Ala Leu Thr			
545	550	555	560
Phe Gln Asp Pro Ser Gly Gly Ser Arg Thr Val Gln Val Thr Leu Gly			
565	570	575	
Lys Ala Glu Gln			
580			

(2) INFORMATION FOR SEQ ID NO:76:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 213 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:76:

Met Asn Asp Gly Lys Arg Ala Val Thr Ser Ala Val Leu Val Val Leu			
1	5	10	15
Gly Ala Cys Leu Ala Leu Trp Leu Ser Gly Cys Ser Ser Pro Lys Pro			
20	25	30	
Asp Ala Glu Glu Gln Gly Val Pro Val Ser Pro Thr Ala Ser Asp Pro			
35	40	45	
Ala Leu Leu Ala Glu Ile Arg Glu Ser Leu Asp Ala Thr Lys Gly Leu			
50	55	60	
Thr Ser Val His Val Ala Val Arg Thr Thr Gly Lys Val Asp Ser Leu			

65	70	75	80
Leu	Gly	Ile	Thr
Ser	Ala	Asp	Val
Asp	Val	Asp	Val
Arg	Ala	Asn	Pro
Asn	Pro	Leu	Ala
	85	90	95
Ala	Lys	Gly	Val
Cys	Thr	Tyr	Asn
Asp	Glu	Gln	Gly
100	105	110	
Val	Gln	Gly	Asp
Asp	Asn	Ile	Ser
Val	Val	Lys	Leu
Phe	Asp	Asp	Trp
Asp	Trp	Ser	Asn
115	120	125	
Leu	Gly	Ser	Ile
Ile	Ser	Glu	Leu
Ser	Thr	Ser	Arg
Arg	Val	Leu	Asp
130	135	140	
Ala	Gly	Val	Thr
Gln	Leu	Leu	Ser
Gly	Val	Thr	Asn
Leu	Gln	Leu	Gly
	145	150	155
Gly	Thr	Glu	Val
Ile	Asp	Gly	Ile
Ser	Thr	Thr	Lys
Thr	Asp	Ile	Thr
165	170	175	
Ile	Pro	Ala	Ser
Ser	Ser	Val	Val
Met	Leu	Asp	Pro
Gly	Ala	Pro	Gly
Ala			Lys
	180	185	190
Arg	Pro	Ala	Thr
Val	Trp	Ile	Ala
		Gln	Asp
		Gly	Ser
		His	His
		Leu	Val
195	200	205	
Arg	Ala	Ser	Ile
Ile	Asp	Leu	Gly
Ser	Gly	Ser	Gly
Ile	Ile	Gly	Ser
210	215	220	
Lys	Trp	Asn	Glu
Asn	Glu	Pro	Val
Asp	Val	Asp	
225	230		

(2) INFORMATION FOR SEC ID NO. 53:

SEQUENCE CHARACTERISTICS

- (A) LENGTH: 66 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 22.

Val	Ile	Asp	Ile	Ile	Gly	Thr	Ser	Pro	Thr	Ser	Trp	Glu	Gln	Ala	Ala			
1																		
														10	15			
Ala	Glu	Ala	Val	Gln	Arg	Ala	Arg	Asp	Ser	Val	Asp	Asp	Ile	Arg	Val			
														20	25	30		
Ala	Arg	Val	Ile	Gln	Gln	Asp	Met	Ala	Val	Asp	Ser	Ala	Gly	Lys	Ile			
															35	40	45	
Thr	Tyr	Arg	Ile	Lys	Leu	Glu	Val	Ser	Phe	Lys	Met	Arg	Pro	Ala	Gln			
															50	55	60	
Pro	Arg																	
															65			

12) INFORMATION FOR SEQ ID NO: 78

(ii) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 69 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(x4) SEQUENCE DESCRIPTION: SEQ ID NO:1a:

Val Pro Pro Ala Pro Pro Leu Pro Pro Leu Pro Pro Ser Pro Ile Ser
1 5 10 15

Cys Ala Ser Pro Pro Ser Pro Pro Leu Pro Pro Ala Pro Pro Val Ala
 20 25 30
 Pro Gly Pro Pro Met Pro Pro Leu Asp Pro Trp Pro Pro Ala Pro Pro
 35 40 45
 Leu Pro Tyr Ser Thr Pro Pro Gly Ala Pro Leu Pro Pro Ser Pro Pro
 50 55 60
 Ser Pro Pro Leu Pro
 65

(2) INFORMATION FOR SEQ ID NO:79:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 355 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:79:

Met Ser Asn Ser Arg Arg Arg Ser Leu Arg Trp Ser Trp Leu Leu Ser
 1 5 10 15
 Val Leu Ala Ala Val Gly Leu Gly Leu Ala Thr Ala Pro Ala Gln Ala
 20 25 30
 Ala Pro Pro Ala Leu Ser Gln Asp Arg Phe Ala Asp Phe Pro Ala Leu
 35 40 45
 Pro Leu Asp Pro Ser Ala Met Val Ala Gln Val Ala Pro Gln Val Val
 50 55 60
 Asn Ile Asn Thr Lys Leu Gly Tyr Asn Asn Ala Val Gly Ala Gly Thr
 65 70 75 80
 Gly Ile Val Ile Asp Pro Asn Gly Val Val Leu Thr Asn Asn His Val
 85 90 95
 Ile Ala Gly Ala Thr Asp Ile Asn Ala Phe Ser Val Gly Ser Gly Gln
 100 105 110
 Thr Tyr Gly Val Asp Val Val Gly Tyr Asp Arg Thr Gln Asp Val Ala
 115 120 125
 Val Leu Gln Leu Arg Gly Ala Gly Leu Pro Ser Ala Ala Ile Gly
 130 135 140
 Gly Gly Val Ala Val Gly Glu Pro Val Val Ala Met Gly Asn Ser Gly
 145 150 155 160
 Gly Gln Gly Gly Thr Pro Arg Ala Val Pro Gly Arg Val Val Ala Leu
 165 170 175
 Gly Gln Thr Val Gln Ala Ser Asp Ser Leu Thr Gly Ala Glu Glu Thr
 180 185 190
 Leu Asn Gly Leu Ile Gln She Asp Ala Ala Ile Gln Pro Gly Asp Ser
 195 200 205
 Gly Gly Pro Val Val Asn Gly Leu Gly Gln Val Val Gly Met Asn Thr
 210 215 220
 Ala Ala Ser Asp Asn Phe Gln Leu Ser Gln Gly Gln Gln Gly Phe Ala
 225 230 235 240
 Ile Pro Ile Gly Gln Ala Met Ala Ile Ala Gly Gln Ile Arg Ser Gly
 245 250 255
 Gly Gly Ser Pro Thr Val His Ile Gly Pro Thr Ala Phe Leu Gly Leu
 260 265 270
 Gly Val Val Asp Asn Asn Gly Asn Gly Ala Arg Val Gln Arg Val Val
 275 280 285

Gly Ser Ala Pro Ala Ala Ser Leu Gly Ile Ser Thr Gly Asp Val Ile
 290 295 300
 Thr Ala Val Asp Gly Ala Pro Ile Asn Ser Ala Thr Ala Met Ala Asp
 305 310 315 320
 Ala Leu Asn Gly His His Pro Gly Asp Val Ile Ser Val Asn Trp Gln
 325 330 335
 Thr Lys Ser Gly Gly Thr Arg Thr Gly Asn Val Thr Leu Ala Glu Gly
 340 345 350
 Pro Pro Ala
 355

(2) INFORMATION FOR SEQ ID NO:80:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 305 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:80:

Ser Pro Lys Pro Asp Ala Glu Glu Gln Gly Val Pro Val Ser Pro Thr
 1 5 10 15
 Ala Ser Asp Pro Ala Leu Leu Ala Glu Ile Arg Gln Ser Leu Asp Ala
 20 25 30
 Thr Lys Gly Leu Thr Ser Val His Val Ala Val Arg Thr Thr Gly Lys
 35 40 45
 Val Asp Ser Leu Leu Gly Ile Thr Ser Ala Asp Val Asp Val Arg Ala
 50 55 60
 Asn Pro Leu Ala Ala Lys Gly Val Cys Thr Tyr Asn Asp Glu Gln Gly
 65 70 75 80
 Val Pro Phe Arg Val Gin Gly Asp Asn Ile Ser Val Lys Leu Phe Asp
 85 90 95
 Asp Trp Ser Asn Leu Gly Ser Ile Ser Gln Leu Ser Thr Ser Arg Val
 100 105 110
 Leu Asp Pro Ala Ala Gly Val Thr Gln Leu Leu Ser Gly Val Thr Asn
 115 120 125
 Leu Gln Ala Gln Gly Thr Glu Val Ile Asp Gly Ile Ser Thr Thr Lys
 130 135 140
 Ile Thr Gly Thr Ile Pro Ala Ser Ser Val Lys Met Leu Asp Pro Gly
 145 150 155 160
 Ala Lys Ser Ala Arg Pro Ala Thr Val Trp Ile Ala Gln Asp Gly Ser
 165 170 175
 His His Leu Val Arg Ala Ser Ile Asp Leu Gly Ser Gly Ser Ile Gln
 180 185 190
 Leu Thr Gln Ser Lys Trp Asn Glu Pro Val Asn Val Asp
 195 200 205

(2) INFORMATION FOR SEQ ID NO:81:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 286 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:81:

Gly Asp Ser Phe Tyr Ala Ala Asp Gln Met Ala Arg Gly Phe Val
 1 5 10 15
 Leu Gly Ala Thr Ala Gly Arg Thr Thr Leu Thr Gly Glu Gly Leu Glu
 20 25 30
 His Ala Asp Gly His Ser Leu Leu Leu Asp Ala Thr Asn Pro Ala Val
 35 40 45
 Val Ala Tyr Asp Pro Ala Phe Ala Tyr Glu Ile Gly Tyr Ile Xaa Glu
 50 55 60
 Ser Gly Leu Ala Arg Met Cys Gly Glu Asn Pro Glu Asn Ile Phe Phe
 65 70 75 80
 Tyr Ile Thr Val Tyr Asn Glu Pro Tyr Val Glu Gln Pro Pro Glu Pro Glu
 85 90 95
 Asn Phe Asp Pro Glu Gly Val Leu Gly Gly Ile Tyr Arg Tyr His Ala
 100 105 110
 Ala Thr Glu Gln Arg Thr Asn Lys Xaa Glu Ile Leu Ala Ser Gly Val
 115 120 125
 Ala Met Pro Ala Ala Leu Arg Ala Ala Glu Met Leu Ala Ala Glu Trp
 130 135 140
 Asp Val Ala Ala Asp Val Trp Ser Val Thr Ser Trp Gly Glu Leu Asn
 145 150 155 160
 Arg Asp Gly Val Val Ile Glu Thr Glu Lys Leu Arg His Pro Asp Arg
 165 170 175
 Pro Ala Gly Val Pro Tyr Val Thr Arg Ala Leu Glu Asn Ala Arg Gly
 180 185 190
 Pro Val Ile Ala Val Ser Asp Trp Met Arg Ala Val Pro Glu Gln Ile
 195 200 205
 Arg Pro Trp Val Pro Gly Thr Tyr Leu Thr Leu Gly Thr Asp Gly Phe
 210 215 220
 Gly Phe Ser Asp Thr Arg Pro Ala Gly Arg Arg Tyr Phe Asn Thr Asp
 225 230 235 240
 Ala Glu Ser Glu Val Gly Arg Gly Phe Gly Arg Gly Trp Pro Gly Arg
 245 250 255
 Arg Val Asn Ile Asp Pro Phe Gly Ala Gly Arg Gly Pro Pro Ala Glu
 260 265 270
 Leu Pro Gly Phe Asp Glu Gly Gly Glu Leu Arg Pro Xaa Lys
 275 280 285

(12) INFORMATION FOR SEQ ID NO:82:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 173 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:82:

Thr Lys Phe His Ala Leu Met Gln Glu Gln Ile His Asn Glu Phe Thr
 1 5 10 15
 Ala Ala Gln Gln Tyr Val Ala Ile Ala Val Tyr Phe Asp Ser Glu Asp
 20 25 30
 Leu Pro Gln Leu Ala Lys His Phe Tyr Ser Ser Gln Ala Val Glu Glu Arg

III

35	40	45
Asn His Ala Met Met Leu Val Gln His Leu Leu Asp Arg Asp Leu Arg		
50	55	60
Val Glu Ile Pro Gly Val Asp Thr Val Arg Asn Gln Phe Asp Arg Pro		
65	70	75
Arg Glu Ala Leu Ala Leu Ala Leu Asp Gin Glu Arg Thr Val Thr Asp		80
85	90	95
Gln Val Gly Arg Leu Thr Ala Val Ala Arg Asp Glu Gly Asp Phe Leu		
100	105	110
Gly Glu Gln Phe Met Gln Trp Phe Leu Gln Glu Gln Ile Glu Glu Val		
115	120	125
Ala Leu Met Ala Thr Leu Val Arg Val Ala Asp Arg Ala Gly Ala Asn		
130	135	140
Leu Phe Glu Leu Gln Asn Phe Val Ala Arg Glu Val Asp Val Ala Pro		
145	150	155
Ala Ala Ser Gly Ala Pro His Ala Ala Gly Gly Arg Leu		160
155		170

(2) INFORMATION FOR SEQ ID NO:83:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 107 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:83:

Arg Ala Asp Glu Arg Lys Asn Thr Thr Met Lys Met Val Lys Ser Ile			
1	5	10	15
Ala Ala Gly Leu Thr Ala Ala Ala Ala Ile Gly Ala Ala Ala Ala Gly			
20	25	30	
Val Thr Ser Ile Met Ala Gly Gly Pro Val Val Tyr Gln Met Gln Pro			
35	40	45	
Val Val Phe Gly Ala Pro Leu Pro Leu Asp Pro Xaa Ser Ala Pro Xaa			
50	55	60	
Val Pro Thr Ala Ala Gln Trp Thr Xaa Leu Leu Asn Xaa Leu Xaa Asp			
65	70	75	80
Pro Asn Val Ser Phe Xaa Asn Lys Gly Ser Leu Val Gln Gly Gly Ile			
85	90	95	
Gly Gly Xaa Glu Gly Xaa Xaa Arg Arg Xaa Gln			
100	105		

(2) INFORMATION FOR SEQ ID NO:84:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 125 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:84:

Val Leu Ser Val Pro Val Gly Asp Gly Phe Trp Xaa Arg Val Val Asn			
1	5	10	15
Pro Leu Gly Gln Pro Ile Asp Gly Arg Gly Asp Val Asp Ser Asp Thr			

20	25	30
Arg Arg Ala Leu Glu Leu Glu Gln Ala Pro Ser Val Val Xaa Arg Gln Gly		
35	40	45
Val Lys Glu Pro Leu Xaa Thr Gly Ile Lys Ala Ile Asp Ala Met Thr		
50	55	60
Pro Ile Gly Arg Gly Gln Arg Gln Leu Ile Ile Gly Asp Arg Lys Thr		
65	70	75
Gly Lys Asn Arg Arg Leu Cys Arg Thr Pro Ser Ser Asn Gln Arg Glu		80
85	90	95
Glu Leu Gly Val Arg Trp Ile Pro Arg Ser Arg Cys Ala Cys Val Tyr		
100	105	110
Val Gly His Arg Ala Arg Arg Gly Thr Tyr His Arg Arg		
115	120	125

(2) INFORMATION FOR SEQ ID NO:85:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 117 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:85:

1	5	10	15
Val Asp Gln Gln Leu Val Thr Arg Val Pro Gln Gly Trp Ser Phe Ala			
30	35	40	45
Gln Ala Ala Ala Val Pro Val Val Phe Leu Thr Ala Trp Tyr Gly Leu			
50	55	60	65
Ala Asp Leu Ala Glu Ile Lys Ala Gly Glu Ser Val Leu Ile His Ala			
70	75	80	85
Gly Thr Gly Gly Val Gly Met Ala Ala Val Gln Leu Ala Arg Gln Trp			
95	100	105	110
Gly Val Gln Val Phe Val Thr Ala Ser Arg Gly Lys Trp Asp Thr Leu			
115	120	125	
Arg Ala Xaa Xaa Phe Asp Asp Xaa Pro Tyr Arg Xaa Phe Pro His Xaa			
135			

(2) INFORMATION FOR SEQ ID NO:86:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 103 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:86:

1	5	10	15
Met Tyr Arg Phe Ala Cys Arg Thr Leu Met Leu Ala Ala Cys Ile Leu			
20	25	30	
Ala Thr Gly Val Ala Gly Leu Gly Val Gly Ala Gln Ser Ala Ala Gln			

Thr Ala Pro Val Pro Asp Tyr Tyr Ttp Cys Pro Gly Gln Pro Phe Asp
 35 40 45
 Pro Ala Trp Gly Pro Asn Trp Asp Pro Tyr Thr Cys His Asp Asp Phe
 50 55 60
 His Arg Asp Ser Asp Gly Pro Asp His Ser Arg Asp Tyr Pro Gly Pro
 65 70 75 80
 Ile Leu Glu Gly Pro Val Leu Asp Asp Pro Gly Ala Ala Pro Pro Pro
 85 90 95
 Pro Ala Ala Gly Gly Gly Ala
 100

(2) INFORMATION FOR SEQ ID NO:87:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 98 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:87:

Val Gln Cys Arg Val Trp Leu Glu Ile Gln Trp Arg Gly Met Leu Gly
 1 5 10 15
 Ala Asp Gln Ala Arg Ala Gly Gly Pro Ala Arg Ile Trp Arg Glu His
 20 25 30
 Ser Met Ala Ala Met Lys Pro Arg Thr Gly Asp Gly Pro Leu Glu Ala
 35 40 45
 Thr Lys Glu Gly Arg Gly Ile Val Met Arg Val Pro Leu Glu Gly Gly
 50 55 60
 Gly Arg Leu Val Val Glu Leu Thr Pro Asp Glu Ala Ala Ala Leu Gly
 65 70 75 80
 Asp Glu Leu Lys Gly Val Thr Ser
 85

(2) INFORMATION FOR SEQ ID NO:88:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 95 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:88:

Thr Asp Ala Ala Thr Leu Ala Gln Glu Ala Gly Asn Phe Glu Arg Ile
 1 5 10 15
 Ser Gly Asp Leu Lys Thr Gln Ile Asp Gln Val Glu Ser Thr Ala Gly
 20 25 30
 Ser Leu Gln Gly Gln Trp Arg Gly Ala Ala Gly Thr Ala Ala Gln Ala
 35 40 45
 Ala Val Val Arg Phe Gln Glu Ala Ala Asn Lys Gln Lys Gln Glu Leu
 50 55 60
 Asp Glu Ile Ser Thr Asn Ile Arg Gln Ala Gly Val Gln Tyr Ser Arg
 65 70 75 80
 Ala Asp Glu Glu Gln Gln Gln Ala Leu Ser Ser Gln Met Gly Phe

86

90

95

(2) INFORMATION FOR SEQ ID NO:89:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 166 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:89:

Met	Thr	Gln	Ser	Gln	Thr	Val	Thr	Val	Asp	Gln	Gln	Glu	Ile	Leu	Asn
1															
													10		15
Arg	Ala	Asn	Glu	Val	Glu	Ala	Pro	Met	Ala	Asp	Pro	Pro	Thr	Asp	Val
								20		25				30	
Pro	Ile	Thr	Pro	Cys	Glu	Leu	Thr	Xaa	Xaa	Lys	Asn	Ala	Ala	Gln	Gln
								35		40				45	
Kaa	Val	Leu	Ser	Ala	Asp	Asn	Met	Arg	Glu	Tyr	Leu	Ala	Ala	Gly	Ala
								50		55				60	
Lys	Glu	Arg	Gln	Arg	Leu	Ala	Thr	Ser	Leu	Arg	Asn	Ala	Ala	Lys	Xaa
								65		70				75	
Tyr	Gly	Glu	Val	Asp	Glu	Glu	Ala	Ala	Thr	Ala	Leu	Asp	Asn	Gly	
														80	
								85		90				95	
Glu	Gly	Thr	Val	Gln	Ala	Glu	Ser	Ala	Gly	Ala	Val	Gly	Gly	Asp	Ser
								100		105				110	
Ser	Ala	Glu	Leu	Thr	Asp	Thr	Pro	Arg	Val	Ala	Thr	Ala	Gly	Glu	Pro
								115		120				125	
Asn	Phe	Met	Asp	Leu	Lys	Glu	Ala	Ala	Arg	Lys	Leu	Glu	Thr	Gly	Asp
								130		135				140	
Gln	Gly	Ala	Ser	Leu	Ala	His	Xaa	Gly	Asp	Gly	Trp	Asn	Thr	Xaa	Thr
								145		150				155	
Leu	Thr	Leu	Gln	Gly	Asp										160
								165							

(2) INFORMATION FOR SEQ ID NO:90:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 5 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:90:

Arg	Ala	Glu	Arg	Met
1				5

(2) INFORMATION FOR SEQ ID NO:91:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 263 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:91:

Val Ala Trp Met Ser Val Thr Ala Gly Gln Ala Glu Leu Thr Ala Ala
 1 5 10 15
 Gln Val Arg Val Ala Ala Ala Tyr Glu Thr Ala Tyr Gly Leu Thr
 20 25 30
 Val Pro Pro Pro Val Ile Ala Glu Asn Arg Ala Glu Leu Met Ile Leu
 35 40 45
 Ile Ala Thr Asn Leu Leu Gly Gin Asn Thr Pro Ala Ile Ala Val Asn
 50 55 60
 Glu Ala Glu Tyr Gly Glu Met Trp Ala Gln Asp Ala Ala Ala Met Phe
 65 70 75 80
 Gly Tyr Ala Ala Thr Ala Thr Ala Thr Ala Thr Leu Leu Pro Phe
 85 90 95
 Glu Glu Ala Pro Glu Met Thr Ser Ala Gly Gly Leu Leu Glu Gin Ala
 100 105 110
 Ala Ala Val Glu Glu Ala Ser Asp Thr Ala Ala Ala Asn Glu Leu Met
 115 120 125
 Asn Asn Val Pro Glu Ala Leu Lys Glu Leu Ala Glu Pro Thr Glu Gly
 130 135 140
 Thr Thr Pro Ser Ser Lys Leu Gly Gly Leu Trp Lys Thr Val Ser Pro
 145 150 155 160
 His Arg Ser Pro Ile Ser Asn Met Val Ser Met Ala Asn Asn His Met
 165 170 175
 Ser Met Thr Asn Ser Gly Val Ser Met Thr Asn Thr Leu Ser Ser Met
 180 185 190
 Leu Lys Gly Phe Ala Pro Ala Ala Ala Ala Gln Ala Val Gln Thr Ala
 195 200 205
 Ala Gln Asn Gly Val Arg Ala Met Ser Ser Leu Gly Ser Ser Leu Gly
 210 215 220
 Ser Ser Gly Leu Gly Gly Gly Val Ala Ala Asn Leu Gly Arg Ala Ala
 225 230 235 240
 Ser Val Arg Tyr Gly His Arg Asp Gly Gly Lys Tyr Ala Xaa Ser Gly
 245 250 255
 Arg Arg Asn Gly Gly Pro Ala
 260

(2) INFORMATION FOR SEQ ID NO:92:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 103 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOROLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:92:

Met Thr Tyr Ser Pro Gly Asn Pro Gly Tyr Pro Gln Ala Gln Pro Ala
 1 5 10 15
 Gly Ser Tyr Gly Gly Val Thr Pro Ser Phe Ala His Ala Asp Glu Gly
 20 25 30
 Ala Ser Lys Leu Pro Met Tyr Leu Asn Ile Ala Val Ala Val Leu Gly
 35 40 45
 Leu Ala Ala Tyr Phe Ala Ser Phe Gly Pro Met Phe Thr Leu Ser Thr

50	55	60	
Glu Leu Gly Gly Asp Gly Ala Val Ser Gly Asp Thr Gly Leu Pro			
65	70	75	80
Val Gly Val Ala Leu Leu Ala Leu Leu Ala Gly Val Val Leu Val			
85	90	95	
Pro Lys Ala Lys Ser His Val Thr Val Val Ala Val Leu Gly Val Leu			
100	105	110	
Gly Val Phe Leu Met Val Ser Ala Thr Phe Asn Lys Pro Ser Ala Tyr			
115	120	125	
Ser Thr Gly Trp Ala Leu Trp Val Val Leu Ala Phe Ile Val Phe Gin			
130	135	140	
Ala Val Ala Ala Val Leu Ala Leu Leu Val Glu Thr Gly Ala Ile Thr			
145	150	155	160
Ala Pro Ala Pro Arg Pro Lys Phe Asp Pro Tyr Gly Glu Tyr Gly Arg			
165	170	175	
Tyr Gly Gln Tyr Gly Gln Tyr Gly Val Gln Pro Gly Gly Tyr Tyr Gly			
180	185	190	
Gln Gln Gly Ala Gln Gln Ala Ala Gly Leu Gln Ser Pro Gly Pro Gln			
195	200	205	
Gln Ser Pro Gln Pro Pro Gly Tyr Gly Ser Gln Tyr Gly Gly Tyr Ser			
210	215	220	
Ser Ser Pro Ser Gln Ser Gly Ser Gly Tyr Thr Ala Gln Pro Pro Ala			
225	230	235	240
Gln Pro Pro Ala Gln Ser Gly Ser Gln Gln Ser His Gln Gly Pro Ser			
245	250	255	
Thr Pro Pro Thr Gly Phe Pro Ser Phe Ser Pro Pro Pro Pro Val Ser			
260	265	270	
Ala Gly Thr Gly Ser Gln Ala Gly Ser Ala Pro Val Asn Tyr Ser Asn			
275	280	285	
Pro Ser Gly Gly Glu Gln Ser Ser Ser Pro Gly Gly Ala Pro Val			
290	295	300	

(2) INFORMATION FOR SWO ID NO. 82

SCIENTIFIC CHARACTERISTICS

- (A) LENGTH: 28 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 20

Gly Cys Gly Glu Thr Asp Ala Ala Thr Leu Ala Gln Glu Ala Gly Asn
 1 5 10 15
 Phe Glu Arg Ile Ser Gly Asp Leu Lys Thr Gln Ile
 20 25

(2) INFORMATION FOR SED ID NO: 84

144 SEQUENCE CHANGES IN GENE

ADVANCED CHARACTERISTICS:
(A) LENGTH: 16 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:94:

Asp	Gln	Vai	Glu	Ser	Thr	Ala	Gly	Ser	Leu	Gln	Gly	Gln	Trp	Arg	Gly
1															
							S								
														10	
															15

(2) INFORMATION FOR SEQ ID NO:95:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 27 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:95:

Gly	Cys	Gly	Ser	Thr	Ala	Gly	Ser	Leu	Gln	Gly	Gln	Trp	Arg	Gly	Ala
1															
							S								10
															15
															20
															25

(2) INFORMATION FOR SEQ ID NO:96:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 27 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:96:

Gly	Cys	Gly	Gly	Thr	Ala	Ala	Gln	Ala	Ala	Val	Val	Arg	Phe	Gln	Glu
1															
							S								10
															15
															20
															25

(2) INFORMATION FOR SEQ ID NO:97:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 27 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:97:

Gly	Cys	Gly	Ala	Asn	Lys	Gln	Lys	Gln	Glu	Leu	Asp	Glu	Ile	Ser	Thr
1															
							S								10
															15
															20
															25

(2) INFORMATION FOR SEQ ID NO:98:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 28 amino acids

- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:98:

```

Gly Cys Gly Ile Arg Gin Ala Gly Val Gln Tyr Ser Arg Ala Asp Glu
1           5           10          15
Glu Gln Gln Gln Ala Leu Ser Ser Gln Met Gly Phe
20          25

```

(2) INFORMATION FOR SEQ ID NO:99:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 507 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:99:

```

ATGAAAGATG TGAATTCGAT CCGCCGAGGT CTGACCCGCGC CGGCTGCAAT CGGCCTCGCT
GGGGCCGGTG TGACTTGGAT CAGGGCTGGC GGCCTGGTGC TATACCGAGAT GCACCGCGTC
GTCTTGGCG CGGGCACTGGC GTGGGACCCG GCATCCGGCCCT CTGACTTGGC GACCCGGCC
CAGTTGGACCA CGGCTCTCAA GATCCCAAGG TGCTTTGGC GACACAAGGC 60
AGCTCTGGCG AGGGCGCGAT CGGGGGCACT GAGGGCGGCA TGCGCGACCA CAACTTGAG
AAGGGCCGCG AGCACCGGGA TGTGGCGCTG TGCTTCAAGC TGACGAAACAT CGACCGGGCG
GCGCCGGGTG CGGCGCACCGC CGAGTTTGGC GTCTCGGTC CGAAGCTCTC GTGGCGGGTC 120
ACGGCGAAAGC TCACGTTCTG GAATCAAGGC GGCTGGATGC TGTCAACCGC ATCGGGCGATC
GAGTTGGCTGC AGGGCGKAGG GAACCTGA 180
                                         240
                                         300
                                         360
                                         420
                                         480
                                         507

```

(3) INFORMATION FOR SEQ ID NO:100:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 158 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:100:

```

Met Lys Met Val Lys Ser Ile Ala Ala Gly Leu Thr Ala Ala Ala Ala
1           5           10          15
Ile Gly Ala Ala Ala Ala Gly Val Thr Ser Ile Met Ala Gly Gly Pro
20          25          30
Val Val Tyr Gln Met Gln Pro Val Val Phe Gly Ala Pro Leu Pro Leu
35          40          45
Asp Pro Ala Ser Ala Pro Asp Val Pro Thr Ala Ala Gln Leu Thr Ser
50          55          60
Leu Leu Asn Ser Leu Ala Asp Pro Asn Val Ser Phe Ala Asn Lys Gly
65          70          75          80
Ser Leu Val Glu Gly Gly Ile Gly Gly Thr Glu Ala Arg Ile Ala Asp
85          90          95
His Lys Leu Lys Lys Ala Ala Glu His Gly Asp Leu Pro Leu Ser Phe

```

100	105	110
Ser Val Thr Asn Ile Gln Pro Ala Ala Gly Ser Ala Thr Ala Asp		
115	120	125
Val Ser Val Ser Gly Pro Lys Leu Ser Ser Pro Val Thr Gln Asn Val		
130	135	140
Thr Phe Val Asn Gln Gly Gly Trp Met Leu Ser Arg Ala Ser Ala Met		
145	150	155
Glu Leu Leu Gln Ala Ala Gly Asn		160
	165	

(2) INFORMATION FOR SEQ ID NO:101:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 500 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:101:

CGTGGUAAU	TGTTTGACCG	TGGGGCGCGG	GTCGCGCTCG	GGAGATCGCG	TGGACGGCGT	60
CATTAACACC	ACCTGCAATT	ACGGGAGGT	AGTAACTCGC	CTCAACGGGA	CGGATCCGGG	120
GGCTGGCCCA	CAGTTCAACG	CCTCACCGGT	GGGGCACTGGC	TATTCGGCA	ATTTCTCTGC	180
CGCACCGCCA	CCTTACGCGG	CTCGCACTGGC	CGCGCAATTG	CAAGCTTGCG	CGGGGGCGGC	240
ACATACATC	GCCTCTTGGC	AGTGGGTTTCG	CGCGCTCTGC	AACAATATT	AAGCCCATGC	300
GGGCCCTCATC	CGGGGACCGG	GCATGCTCGC	CGGGGCTTGC	CCAGATGGCC	CGCGCTCTCA	360
ACGGGCGGAA	TCTTGACCC	CGGCAATGTC	CGCGGGGCTA	CGCGAGATTG	CCCCGCTCT	420
CAACGGGCCG	CATCGCTGC	CGAATTCCTG	CGGGGGGGG	GATCCATAG	TTCTAGAGCG	480
GGCGGACCGG	CGCTGGAGCT					500

(2) INFORMATION FOR SEQ ID NO:102:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 96 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:102:

Val Ala Met Ser Leu Thr Val Gly Ala Gly Val Ala Ser Ala Asp Pro					
1	5	10	15		
Val Asp Ala Val Ile Asn Thr Thr Cys Asn Tyr Gly Gln Val Val Ala					
20	25	30	35		
Ala Leu Asn Ala Thr Asp Pro Gly Ala Ala Ala Gln Phe Asn Ala Ser					
35	40	45			
Pro Val Ala Gln Ser Tyr Leu Arg Asn Phe Leu Ala Ala Pro Pro Pro					
50	55	60			
Gln Arg Ala Ala Met Ala Ala Gln Leu Gln Ala Val Pro Gly Ala Ala					
65	70	75	80		
Gln Tyr Ile Gly Leu Val Glu Ser Val Ala Gly Ser Cys Asn Asn Tyr					
85	90	95			

(2) INFORMATION FOR SEQ ID NO:103:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 154 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:103:

ATGAGAGAGC AGGAGTGGAA TTTCGGCGGT ATCCAGGGCG CGGCAAGCGC AATCCAGGGA	60
AAATGTCACGT CCATTCACTT CCTCTCTTGAC GAGGGAAAGC AGTCCCTGAC CAAAGCTCGCA	120
GGCGGCTCTGGG GCGGTAGCGG TTCCGAAAGCG TACCG	154

(2) INFORMATION FOR SEQ ID NO:104:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 51 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:104:

Met Thr Gln Gln Trp Asn Phe Ala Gly Ile Glu Ala Ala Ala Ser			
1	5	10	15
Ala Ile Gln Gly Asn Val Thr Ser Ile His Ser Leu Leu Asp Glu Gly			
20	25	30	
Lys Gln Ser Leu Thr Lys Leu Ala Ala Ala Trp Gly Gly Ser Gly Ser			
35	40	45	
Glu Ala Tyr			
50			

(2) INFORMATION FOR SEQ ID NO:105:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 282 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:105:

GGTCCGGCGA CCTCCAGGTG ACTATGAAAG TCGGCTTCGG NCTGGAGGAT TCTGTGAACT	60
TCAAGCGCGG CGGATAACTG AGGTGCACTA TTAAAGCGACT TTTCGAGAAC ATCTGTGACGC	120
GCTCGAAACG CGGACACGCC GACGGTGGGT CGCGNCGAGGC GCTGNCCTGA AAATCCCTGA	180
GACAAATTGCG CGGGGGCGCC TACAGGAAAG TCGGTGCTGA ATTGCGCGNG TATCTGTGCG	240
ACCTGTGTGG TCTNAGCTG GACGAAGCGG TGCTCGACUT CG	282

(2) INFORMATION FOR SEQ ID NO:106:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 3098 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:106:

GATCCTACCC	GTGCGAGTGC	TGGGGCCTTT	TGAGGATGGA	GTGCACTGT	CITTCCTGAT	60
GGCATACCC	GAGATGTTGG	CGCGCGCGGC	TGACACCTTG	CAGAGCATCG	GTGCTACAC	120
TGTGGCTAAGC	AATGCCGCTG	CGGGCGCCCT	GACGACTGGG	GTGGTGCCTC	CGCGTCCGGA	180
TGAGGTCTCG	CGCGTGAATG	CGGGCGACTT	GGGGCGATGT	ATCAGTCGCT		240
GGAGCGCTGG	CGCTCTCGGA	TTCATGACCA	CTTCGTTGGC	ACCGTTGCGA	CGAGCGCCAG	300
CTCGTATGCG	GCCACTGAAG	TGGCAATGC	GGCGCGGCC	AGCTAAGCGA	GGACAGTCG	360
GCACGAGAAA	CCACGAGAAA	TAGGGACACG	TAATGGTGGG	TTTCGGGGCG	TTACCGCCGG	420
AGATCRACTC	CGCGGAGGAT	TACGCGCGC	CGGGTTCCGG	CTCGCTGCTG	GGCGCGCTTC	480
AGATGTGGGA	CACCGTGGCG	AGTGACCTGT	TTTCGGCCCG	GTGCGCGTT	CACTCGCTGG	540
TCTGGGGTCT	GGCGGTGGG	TCTGAGTAG	CTTCGTCGGC	GGGTCGATG	GTGCGCGCGG	600
CTTCGCGCTA	TCTGGCGTGG	ATGAGCGTCA	CGGGGGGCGA	GGCGAGCTG	ACCGCGCGC	660
AGGTGCGGT	TGCTGCGCGG	GGTAAAGAGA	CGCGTATGG	GTGACGCTG	CGCGCGCCGG	720
TGATCSCCGA	GAACCTGTG	GAATGATGA	TTCTGATAGC	GGACCAACCTC	TTGGGGCRAA	780
ACACCCCGGC	GATCGCGCTC	AACGAGGCGG	AATACCGCGA	GATGTTGGCC	CAAGACGCCG	840
CGCGCATGTT	TGGTACGGCT	GGCGCGACCG	CGACCGCGAC	GGCGACGCTG	CTGCGCTTCG	900
AGGAGGCCCG	GGGAGTACCC	AGCGCGGCG	GGCTCTCGA	CGAGGCCCGC	CGGATCGAGG	960
AGGCGTCTGA	CGCGCGCGG	CGCGAACCGT	TGATGACCA	TGTGCGCGAG	GGGCTCGAAC	1020
AGCTGCGCCA	GGCCACCGCG	GGCGCCACGC	CTTCCTCCAA	GTGCGTGGC	CTGTCGAAAGA	1080
CGCTCTCGCC	CGATCGCTG	CGATGACGA	ACATGCTGTC	GATGCCAAC	AACCGCATGT	1140
CGATGACCAA	CTCGGGTGTG	TGATGACCA	ACACCTTGAG	CTCGATGTTG	AAGGGCTTTC	1200
CTCGGGCGGC	GGCCCGCCCG	GGCGTGCAAA	GGCGCGCGCA	AAACGGGGTC	CGGGCGATGA	1260
GTGCGCTGGG	CGCTCGCGT	GGTTCTTCGG	CTCTGGCGG	TGGGGCGCC	GGCAACTTGC	1320
GTGCGCGCCG	CTCGCTCGGT	TGTTTCTCGG	TGCGCGCGC	CTGGGGCGCG	GGCAACCGAG	1380
CGATCAACCC	GGCGCGCGCG	GGCGTGCCTG	TGACCAAGCT	GACCAAGCGG	GGGGAAGAGA	1440
GGCCCGGGCA	GATGCTGGC	GGGCTCCCG	TGGGGCAGAT	GGCGCGCCAGG	GGCGCGTGTG	1500
GGCTGAGTGC	TGTCTGCGT	GTTCGCGCGC	GACCGTATGT	GATGCCCAT	TCTCGCGCG	1560
CGCGCTAGGA	GGGGGGCGCG	AGACTGCTG	TATTTGACCA	GTGATCGCG	GTTCGCGTGT	1620

TTCGCGGCC	GGCTATGACA	ACAGTCATG	TGCATGACAA	GTTACAGGTA	TTAGGTGAG	1680
GTTCARCAAG	GAGCACAGGCA	ACATGGCCTC	ACUTTTTATG	ACGGATCCGC	ACCGATGCG	1740
GGACATGCGG	GGCGCGTTTG	AGGTGACACG	CCAGACGCGG	GAGGACGAGG	CTGGCGGGAT	1800
GTGGGGCGTC	GGCGCAAAAC	TTTCGGTGC	GGGCTGGAGT	GGCATGGCG	AGGGGACCTC	1860
GTCTAGACACC	ATGGCCGAGA	TGAATCGAGC	GTTCGCAAC	ATCGTGAACA	TGCTGCACGG	1920
GGTGGCGTAC	GGGCTGGTTC	GGGAGCGCCAA	CAACTACGAG	CAGCAAGAGC	AGGGCTCCCA	1980
GCAGATCTTC	AGCAGCTAAC	GTCAAGGCGT	GGAGGACAAAT	ACTTTTACAA	GGGAAAGGAGA	2040
ACAGGTTGCA	TGACCATCAA	CTATCAATTG	GGGGATGTCG	ACGCTCACCG	GGCGCTGATC	2100
GGCGCTCAGG	CCGGGTTGCT	GGAGGCGAG	CTTCAGGCGA	TGATTCCTGA	TGTTGTTGACC	2160
GGCAGTGA	ACTTTGGGGCGG	CCCGCGTTGC	GGCGCGTCCG	AGGGGTTCT	TACCCAGTTG	2220
GGCGCGTAACT	TCAGCGTGT	CTAGGAGCG	GGCAACGCC	ACGGGGCAGAA	GGTGCGAGGT	2280
GGCGCGAACAA	ACATGGCGCA	AACCGACAGC	GGCTTGCGCT	CCAGCTGGCG	CTGACACCCAG	2340
GGCGAGGCCA	GGGAGCTGGT	GTACGAGTGA	AGTTCTCGCG	GTGATCCCTTC	GGGTGGCGAT	2400
CTTAACTGGTC	AGTGTCTGGG	TGTTGCTGCT	TTGCTGCTTG	GGGGGTTCTT	GGGTGCTGGT	2460
CACTGCTGCT	GGGGCTGGGG	TGAGGACCTC	GAAGGCCAGG	TAGGCGCTC	CTTGCGATCCA	2520
TTGCTGCTGT	TGTTCGCGCA	GGACGGCTGC	GAAGGAGCGG	ATGATCGGG	GGCGCGTCCGG	2580
GAAGATGCCA	ACGACGCTGG	TTGCGCGCTG	TACCTCTGG	TTGAGGCGTT	CTGGGGGTTT	2640
GTGGGACCG	ATTGGCGCGC	AGATCTGCTT	GGGGAAAGGC	GTGAACGCCA	GGAGGCTGGT	2700
GGGGCGCTGC	TGGAGGTTGT	GGGCCACCCG	GGGGAGTTTG	TCGGTCAGAG	CTGCGAGTAC	2760
CCGATCATRT	TGGGACCAA	CTGATTCGGC	GTGGGCGCTG	TGCTAGATGG	AGTCGAGCAG	2820
GGTGGCGACCC	CAAGGCCAGG	AGGCGCTCGG	GGTGGCTGCC	ATCGAGATGG	CTGGCTAGTG	2880
GTTTCTGCGAG	GGCTGCCAGG	GGGTGCGGG	CAAGGCGGG	CCGATCGGG	CCACCGAGCC	2940
GGCGTGGCGC	TGGCTGCTGA	CCAGCGCGAC	GGCGGACAGG	GGCGGGCGA	CCAGGTGCGC	3000
GAAGAAACCC	AGCCAGCCGG	CTTCTGCTC	GGCGGAGGTC	ACCTGGATGC	CGAGGATC	3068

(2) INFORMATION FOR SEQ ID NO:107:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 391 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS: single

(D) TOPOLOGY: It means

(seq) SEQUENCE DESCRIPTION: SEQ ID NO: 10

Met Val Asp Phe Gly Ala Leu Pro Pro Glu Ile Asn Ser Ala Arg Met
 1 5 10 15
 Tyr Ala Gly Pro Gly Ser Ala Ser Leu Val Ala Ala Ala Gln Met Trp
 20 25 30
 Asp Ser Val Ala Ser Asp Leu Phe Ser Ala Ala Ser Ala Phe Gln Ser
 35 40 45
 Val Val Trp Gly Leu Thr Val Gly Ser Trp Ile Gly Ser Ser Ala Gly
 50 55 60
 Leu Met Val Ala Ala Ala Ser Pro Tyr Val Ala Trp Met Ser Val Thr
 65 70 75 80
 Ala Gly Gln Ala Glu Leu Thr Ala Ala Gln Val Arg Val Ala Ala Ala
 85 90 95
 Ala Tyr Glu Thr Ala Tyr Gly Leu Thr Val Pro Pro Pro Val Ile Ala
 100 105 110
 Glu Asn Arg Ala Glu Leu Met Ile Leu Ile Ala Thr Asn Leu Leu Gly
 115 120 125
 Glu Asn Thr Pro Ala Ile Ala Val Asn Glu Ala Glu Tyr Gly Glu Met
 130 135 140
 Tyr Ala Gln Asp Ala Ala Ala Met Phe Gly Tyr Ala Ala Ala Thr Ala
 145 160 155 160
 Thr Ala Thr Ala Thr Leu Leu Pro Phe Glu Glu Ala Pro Glu Met Thr
 165 170 175
 Ser Ala Gly Gly Leu Leu Glu Gln Ala Ala Ala Val Glu Glu Ala Ser
 180 185 190
 Ser Thr Ala Ala Ala Asn Gln Leu Met Asn Asn Val Pro Gln Ala Leu
 195 200 205
 Asn Gln Leu Ala Gln Pro Thr Gln Gly Thr Thr Pro Ser Ser Lys Leu
 210 215 220
 Lys Gly Leu Trp Lys Thr Val Ser Pro His Arg Ser Pro Ile Ser Asn
 225 230 235 240
 Val Ser Met Ala Asn Asn His Met Ser Met Thr Asn Ser Gly Val
 245 250 255
 Met Thr Asn Thr Leu Ser Ser Met Leu Lys Gly Phe Ala Pro Ala
 260 265 270

Ala Ala Ala Gln Ala Val Gln Thr Ala Ala Gln Asn Gly Val Arg Ala
 275 280 285
 Met Ser Ser Leu Gly Ser Ser Leu Gly Ser Ser Gly Leu Gly Gly Gly
 290 295 300
 Val Ala Ala Asn Leu Gly Arg Ala Ala Ser Val Gly Ser Leu Ser Val
 305 310 315 320
 Pro Gln Ala Trp Ala Ala Asn Gln Ala Val Thr Pro Ala Ala Arg
 325 330 335
 Ala Leu Pro Leu Thr Ser Leu Thr Ser Ala Ala Glu Arg Gly Pro Gly
 340 345 350
 Gln Met Leu Gly Gly Leu Pro Val Gly Gln Met Gly Ala Arg Ala Gly
 355 360 365
 Gly Gly Leu Ser Gly Val Leu Arg Val Pro Pro Arg Pro Tyr Val Met
 370 375 380
 Pro His Ser Pro Ala Ala Gly
 385 390

(2) INFORMATION FOR SEQ ID NO:108:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1728 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: linear

(ii) SEQUENCE DESCRIPTION: SEQ ID NO:108:

GAGTCAGCA	CCGCCCTGTC	AGGGCTGGAG	CCTGGTCGGT	TTTGATCTGC	GCTCAGGGTG	60
ACGGTCCCTCG	GGCTGTGCGCC	GGCTGGGATG	CAGACTCGAT	GCCGCTCTTT	ASTGCACTA	120
ATTTCCTTGA	AGTGGCTCGG	AGGTATAGGA	CTTCAGCATT	GGTTAAATGA	GGCTTCACCC	180
GGTGTGGGG	TCTGTTGGC	GGGACCGCTC	GTCAACCAACG	CTTGGCCTGCG	GGCCCAAGCG	240
GGGGATCAGA	TCGGTGTACT	ACCAATCAAT	CTTGAGCTCC	GGGGCCGATG	CTCGGGCTAA	300
ATGAGGAGGA	GGAGCGGTGT	CTTTCAGCTGC	GGAAACGGAG	ATGTTGGCGG	CCCGCGCTGG	360
CGRAGCTTCGT	TCTCTGGGGG	CAACGCTGAA	GGCTAGCAAT	GGCGCCGGAG	CCGTGCGGAC	420
GACTGGGCGT	GTGCCCCCGG	CTGGCCGACGA	GGTGTGGCTG	CTGCTTGGCA	CAACATTCCG	480
TACCGCATCG	GGGAGCTATC	AGACGGCGAG	GGCGAAGGGC	GGGGTGAATCC	ATGAGGCAATT	540
TGTGACCAAG	CTGGCCACTA	GGCGTAGTTT	ATATGCGGAC	ACCGAGGGCG	CCAAACUCTGT	600
GCTCACCGGC	TAGCTGACCT	GACGGTATTTC	GAGCGGAAGG	ATTATCGAAG	TGGTGGATT	660

CCGGCCCTTA	CCACCGGAGA	TCAACTCCGC	GAGGATGTAC	GGGGGCCCCG	GTTCGGCCCTC	720
GCTGGTGGCC	GGGGCGAAGA	TGTGGACAG	CGTGGCGAGT	GAACCTTTTT	CGGCGCGTC	780
GGCGTTTCAG	TGGTGTGCT	GGGCTCTGAC	GGTGGGTTGC	TGGTAGAGTT	CGTGGCGCG	840
TCTGATGGG	GGGGCGGCCT	CGCGTATGT	GGCGTGGATG	AGCGTCACCG	CGGGCGACCG	900
CGCGCTAACG	GGGGCCAGG	TGGGGTTTC	TGGGGCGGCC	TACGAGACAG	CCTATAGGCT	960
GACGGTGCCTC	GGGGCGGTGA	TGGCCGAGAA	CCGTACCGAA	CTGATGACGC	TGACCGCGAC	1020
CAACCTTCTTG	GGGCAAAACG	CGCCCGCGAT	CGAGGCGAAAT	CAGGCCCGAT	ACAGCCAGAT	1080
GTGGGGCCAA	GGGGCGGAGG	CGATGTATGG	CTACGCGCCG	ACGGCGGGGA	CGGGCGACCGA	1140
GGGGTGTGCTG	GGGGTGTGAGG	ACGGCGCGAT	GTACGACGAC	CGGGCGGGGC	TGTTTGAACA	1200
GGCCGCTGGCG	GTGGAGGAGG	CGATCGACAC	CGGGCGGGCG	AACCGATGAA	TGAACATGT	1260
GGCCCCAAGCG	CTGCAACACG	TGGCCCGAGC	ACGGCGAGGC	GTGGTACCTT	CTTCCGAGCT	1320
GGGTGGGCTG	TGGACGCGCG	TGTGGGGCGA	TGTGTGCGCG	CTGAGCAACG	TCACTTGTAT	1380
AGCCAAACAC	CACTATGCGA	TGATGGCGAC	GGGTGTGTGC	ATGACCAAAAC	CTTGCACTC	1440
GATTTTGAAG	GGCTTAGCTC	GGGGGGGGGC	TCAGGCCGTG	GAACCGCG	CGGAAAAACGG	1500
GGTGTGGCGG	ATGAGGTCTC	TGGCGAGCTA	GTGGGGTTTG	TGTGTGGGTT	CTTCGGGCT	1560
GGGGGCTGGG	GTGGCCCGCA	ACTTGGGCTG	GGGGGCGCTG	GTGGGGTTGT	TGTGGGTGCC	1620
GGCAGGCGATG	GGGGCGCGCA	ACCTAGCGGT	CAACCGCGCT	GGGGGGGGCG	TGCCCCGTGAC	1680
CAACCTTCTTG	AGGGGGCGCCG	AAACCGCGCTC	GGGACACATG	CTGGG		1725

(2) INFORMATION FOR SEO IN NO. 326.

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 339 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: [seq_1000](#)

Val Val Asp Phe GLY Ala Leu Pro Pro Glu Ile Asn Ser Ala Arg Met
1 8 10 15

Tyr Ala Gly Pro Gly Ser Ala Ser Leu Val Ala Ala Ala Lys Met Trp
20 25 30

Asp Ser Val Ala Ser Asp Leu Phe Ser Ala Ala Ser Ala Phe Gln Ser
 35 40 45

Val Val Trp Gly Leu Thr Val Gly Ser Trp Ile Gly Ser Ser Ala Gly
 50 55 50
 Leu Met Ala Ala Ala Ser Pro Tyr Val Ala Trp Met Ser Val Thr
 65 70 75 80
 Ala Gly Gln Ala Gln Leu Thr Ala Ala Gln Val Arg Val Ala Ala Ala
 85 90 95
 Ala Tyr Gln Thr Ala Tyr Arg Leu Thr Val Pro Pro Pro Val Ile Ala
 100 105 110
 Glu Asn Arg Thr Glu Leu Met Thr Leu Thr Ala Thr Asn Leu Leu Gly
 115 120 125
 Gln Asn Thr Pro Ala Ile Glu Ala Asn Gln Ala Ala Tyr Ser Gln Met
 130 135 140
 Trp Gly Gln Asp Ala Glu Ala Met Tyr Gly Tyr Ala Ala Thr Ala Ala
 145 150 155 160
 Thr Ala Thr Glu Ala Leu Leu Pro Phe Glu Asp Ala Pro Leu Ile Thr
 165 170 175
 Asp Pro Gly Gly Leu Leu Glu Gln Ala Val Ala Val Glu Glu Ala Ile
 180 185 190
 Asp Thr Ala Ala Ala Asn Gln Leu Met Asn Asn Val Pro Gln Ala Leu
 195 200 205
 Gln Gln Leu Ala Gln Pro Ala Gln Gly Val Val Pro Ser Ser Lys Leu
 210 215 220
 Gly Gly Leu Trp Thr Ala Val Ser Pro His Leu Ser Pro Leu Ser Asn
 225 230 235 240
 Val Ser Ser Ile Ala Asn Asn His Met Ser Met Met Gly Thr Gly Val
 245 250 255
 Ser Met Thr Asn Thr Leu His Ser Met Leu Lys Gly Leu Ala Pro Ala
 260 265 270
 Ala Ala Gln Ala Val Glu Thr Ala Ala Glu Asn Gly Val Trp Ala Met
 275 280 285
 Ser Ser Leu Gly Ser Gln Leu Gly Ser Ser Leu Gly Ser Ser Gly Leu
 290 295 300
 Gly Ala Gly Val Ala Ala Asn Leu Gly Arg Ala Ala Ser Val Gly Ser
 305 310 315 320
 Leu Ser Val Pro Pro Ala Trp Ala Ala Asn Gln Ala Val Thr Pro
 325 330 335

Ala Ala Arg Ala Leu Pro Leu Thr Ser Leu Thr Ser Ala Ala Gln Thr
 340 345 350

Ala Pro Gly His Met Leu Gly
 355

(2) INFORMATION FOR SEQ ID NO:110:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 3027 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:110:

AGTTTCAGTCG AGATATGATAC TGACCGGGCTG TATCCACGAT GGCTGAGACA ACCGAACAC	60
CGTGGACGC GGGGACATCG CAAACCGAAG CGATUGCGTT GGCTGCGGAA CGCGAAGCG	120
CGGAGGCGAA AGGCTGGCC CGCCCGGCGC GGGCGCGTC CGCTGGCGCC CGTGTGAAAGC	180
GTGAGGCGCT CGCGATGGCC CGACCGCGAGG ACCGAAGACGT CGCGCGAGGAT ATGCAGACTG	240
GGGAGACCGT GGAACTATGAC CGACATATGAG CGACTATGAG GCGCGAGACCC AGGAGAGCGC	300
ACGGTGGCA TGTCTGGAC CGCGGTTGCC CGTGGCGTTA CGAGACTGT CGACGATTGCG	360
CATGGCGCT CGATGCTCA TGTCTGCGG CGTCACCGGG CGTCAGGGAT ACATTGTGTC	420
CGAACACCAT GAGGCCACCG AACCGGAGCA CGCCCGCGCC CGCTTGCGC CGCGACCCAA	480
CGAAGGTTT ATCAACATCA CGTCTGGCG CGAAAGAAG ACCTTGCGCG CGCGACCGAA	540
TGTGATGCGAC AGCTCCACCG CGCGATTGCG CGATGACTTC CGACCGGGCG CGACCGGATT	600
CGACCAAGGT CGCGAACAGT CGAAACTGT CGCTGAAGGC CGCTTGACG CGACGCGGT	660
CGAACTGTCA AACGACATT CGCGCTGTGT CGCTGTGCG CGAGCTTCAC CGGTGACCAA	720
TTCGCGTGG CGCGAAAGACG AACCGCTGCG CGCGCGCTG AAAGTGACCG TGACCGAAAGA	780
GGGGGGACAGC TACAAGATGT CGAAAGTTGA CGTGTGACCG TGACCGAAGA CGTACCGGAC	840
GTCAACACCG AACCACTGA CGCCGACCGA CGCGCTGCGA CGACTCTGAC CGCAGGGCAA	900
CGCGGTGATT CGCGACCGA CGCGATTGAC ACCGACTCTG CGACCGAATC TACCGGGCGAG	960
AAGGTTGAGC CGCACCGTGA CGTGTGGCGA ATGCAGGTTA CGTGTGAAACC CGTTCGGGTG	1020
ATTCGATCC CGTCATGTT GATGCTGCG CGCGGACCG GATGGCTATA CGTGTGAGCAA	1080
TACGACCGA TGCGCGACG GATGCTGCG CGCGCGTGC CGCGTGTGCC CGGGCGTGTG	1140
ACGGGACAAAT CGCGCTGTG TGTATTCACG CGACCGTGC ACCGAGACTT CGCTACCGCC	1200

